

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 13:20:20 ; Search time 232 Seconds
(without alignments)
268.620 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYADGSANMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-DB=/cgm2_1/USPTO_spool/US10751235/runat_12012006_120916_2581/app_query.fasta.1.263
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNIT8=bits -START=1 -END=1 -MAIRIX=blobum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10751235 @CGN 1 1 122 @runat_12012006_120916_2581
-NCPUs=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FEALOP=6 -FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgm2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgm2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgm2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgm2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq.*
- 5: /cgm2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgm2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgm2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	17.4	2118	US-11-136-527-3214	Sequence 3214, Ap
2	67.5	17.2	3198	US-10-750-185-26357	Sequence 26357, A
3	67.5	17.2	3198	US-10-750-623-26357	Sequence 26357, A
4	64.5	16.4	3917	US-10-750-185-31952	Sequence 31952, A
5	64.5	16.4	3917	US-10-750-623-31952	Sequence 31952, A
6	64	16.3	804	US-10-793-626-1611	Sequence 1611, Ap
7	64	16.3	2918	US-10-793-626-3932	Sequence 3932, Ap
8	64	16.3	3660	US-10-793-626-4116	Sequence 4116, Ap

C	9	62	15.8	76329	7	US-11-150-888-16	Sequence 16, Appl
	10	61.5	15.6	2091	7	US-11-052-554A-484	Sequence 484, App
	11	61.5	15.6	2091	7	US-11-052-554A-485	Sequence 485, App
C	12	61.5	15.6	191797	7	US-11-121-086-13	Sequence 13, Appl
	13	61	15.5	2858	7	US-11-128-061-473	Sequence 473, App
	14	61	15.5	11131	6	US-10-240-708-27	Sequence 27, Appl
	15	60.5	15.4	1494	6	US-10-793-626-1623	Sequence 1623, Ap
	16	60.5	15.4	2005	7	US-11-136-527-3439	Sequence 3439, Ap
C	17	60.5	15.4	3721	6	US-10-793-626-3966	Sequence 3966, App
	18	59	15.0	1547	7	US-11-000-463-171	Sequence 171, App
	19	59	15.0	1956	7	US-11-052-554A-495	Sequence 495, App
C	20	59	15.0	2782	7	US-11-128-061-952	Sequence 952, App
	21	58.5	14.9	1016	6	US-10-750-185-42840	Sequence 42840, A
	22	58.5	14.9	1016	6	US-10-750-623-42840	Sequence 42840, A
	23	58.5	14.9	3027	6	US-10-467-657-185	Sequence 185, App
	24	58.5	14.9	3027	6	US-10-467-657-4555	Sequence 4555, App
C	25	58.5	14.9	137000	6	US-10-515-538-11	Sequence 11, Appl
	26	58.5	14.9	158692	7	US-11-121-086-30	Sequence 30, Appl
	27	58.5	14.9	1125000	6	US-10-995-581-13286	Sequence 13286, A
C	28	58	14.8	1260	6	US-10-750-185-52875	Sequence 52875, A
	29	58	14.8	1260	6	US-10-750-623-52875	Sequence 52875, A
C	30	57.5	14.6	879	6	US-10-750-185-30691	Sequence 30691, A
	31	57.5	14.6	879	6	US-10-750-623-30691	Sequence 30691, A
C	32	57.5	14.6	1987	6	US-10-750-185-49778	Sequence 49778, A
	33	57.5	14.6	1987	6	US-10-750-623-49778	Sequence 49778, A
	34	57.5	14.6	3590	7	US-11-136-527-3160	Sequence 3160, Ap
	35	57	14.5	201	6	US-10-995-561-42931	Sequence 42931, A
C	36	57	14.5	2616	6	US-10-995-561-467	Sequence 467, App
	37	57	14.5	2787	6	US-10-995-561-469	Sequence 469, App
C	38	57	14.5	2878	6	US-10-995-561-468	Sequence 468, App
	39	57	14.5	3223	6	US-10-750-185-46677	Sequence 46677, A
C	40	57	14.5	3223	6	US-10-750-623-46677	Sequence 46677, A
	41	56.5	14.4	1347	7	US-11-055-822-909	Sequence 909, App
C	42	56.5	14.4	2233	6	US-10-750-185-43888	Sequence 43888, A
	43	56.5	14.4	2233	6	US-10-750-623-43888	Sequence 43888, A
C	44	56.5	14.4	3138	6	US-10-750-185-45175	Sequence 45175, A
	45	56.5	14.4	3138	6	US-10-750-623-45175	Sequence 45175, A

ALIGNMENTS

RESULT 1

US-11-136-527-3214
; Sequence 3214, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3214
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3214

Alignment Scores:			
Pred. No.:	11.8	Length:	2118
Score:	68.50	Matches:	20
Percent Similarity:	46.15%	Conservative:	10
Best Local Similarity:	30.77%	Mismatches:	30
Query Match:	17.43%	Indels:	5
DB:	7	Gaps:	2

US-10-751-235-1 (1-77) x US-11-136-527-3214 (1-2118)

QY 5 AlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspVal 24

Db	769	TCGGAGGAAGGCAAGCCAC	CAGCATGTAAGACATCTTTGGGGCTACACGATGGATGTG	828
		::: :::	::: :::	
Qy	25	IleGlyLeuSerLeuPheAsnTyrAnPheAspSerLeuThr	Thr---AspSerProVal	43
		:::	:::	
Db	829	ATCACAGCCACCTCATTTGGAGTGAATGTGATTCCTTAA	CAACCCACAGGAGCCCTTTT	888
		:::	:::	
Qy	44	IleGluAlaValTyrThrAlaLeuLys-----GluAlaGluLeuArgSerThr	59	
		:::	:::	
Db	889	GTGAAATAAGTCAAGAAGCTCTTAAATTTGATATCTTTGATC	CATTTGTTCTCTCAGTG	948
		:::	:::	
Qy	60	AspLeuLeuProTyr	64	
Db	949	ACACTTTTTCATTTC	963	

```

RESULT 2
US-10-750-185-26357/c
; Sequence 26357, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26357
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: Bovine 19866881124994
US-10-750-185-26357

Alignment Scores:
Pred. No.:          30 3      Length:       3198
Score:              67.50     Matches:      17
Percent Similarity: 42.59%    Conservative:   6
Best Local Similarity: 31.48% Mismatches:      28
Query Match:        17.18% Indels:         3
DB:                  6       Gaps:         1

US-10-751-235-1 (1-77) x US-10-750-185-26357 (1-3198)

Qy      25 IleGlyLeuSerLeuPheAsnTyAsnPhaAppSerLeuThrThraSpSerProValIle 44
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2235 ATACCAAGCTGAAGTTCCTTAATATGTCGGATAACGGGAAACGTGCCTCTCCTAGATG 2176

Qy      45 GluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu----- 61
      ::: :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2175 AGCCCTTCAGTTGTGGCGCTCTCCATCACCTCCCTTTCCCCAACACACTGAAAAACAGTT 2116

Qy      62 LeuProTyTrpLysIleAspAlaLeuCysLysIleValPro 75
      ::: :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2115 CTTCCTCTATGGCTGTGTTCTCTCAAGCTGCAAAATTATTTCACA 2074

```

```

; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26357
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: Bovine 19866881124994
US-10-750-623-26357

Alignment Scores:
Pred. No.: 30.3 Length: 3198
Score: 67.50 Matches: 17
Percent Similarity: 42.59% Conservative: 6
Best Local Similarity: 31.48% Mismatches: 28
Query Match: 17.18% Indels: 3
DB: 6 Gaps: 1

US-10-751-2335-1 (1-77) x US-10-750-623-26357 (1-3198)

Qy 25 IleGlyLeuSerLeuPheAenTyAenPheAspSerLeuThrThrAspSerProValIle 44
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2235 ATACCAAGCTGAGTCTCTTAATATGTGCGGATACGGGAACGTGACCTCTCCTAGATG 2176

Qy 45 GluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu----- 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2175 AGCCTTCAGTTTGTGGCTCTCCTCATCACCTCCCTTCCCAACACACTGAAACAGTT 2116

Qy 62 LeuProTyTrpLysIleAspAlaLeuCysLysIleValPro 75
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2115 CTTCCTCTATGGACTGCTCTTCCTCAAGCTGCAGAAATTATTTC 2074

```

```

RESULT 4
US-10-750-185-31952/c
; Sequence 31952, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31952
; LENGTH: 3917
; TYPE: DNA
; ORGANISM: Bovine 1986680527023
US-10-750-185-31952

Alignment Scores:
Pred. No.: 119
Score: 64.50
Percent Similarity: 45.68%
Best local Similarity: 27.16%
Query Match: 16.41%
DB: 6

Length: 3917
Matches: 22
Conservative: 15
Mismatches: 29
Indels: 15
Gaps: 4

US-10-751-235-1 (1-77) x US-10-750-185-31952 (1-31917)

```



```

; LENGTH: 2918
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description

```



```
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 473
; LENGTH: 2858
; TYPE: DNA
; ORGANISM: Cricetinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (129)..(191)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (366)..(400)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)..(501)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1020)..(1034)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-473
Alignment Scores:
Pred. No.: 266 Length: 2858
Score: 61.00 Matches: 25
Percent Similarity: 43.75% Conservative: 10
Best Local Similarity: 31.25% Mismatches: 37
Query Match: 15.52% Indels: 8
DB: 7 Gaps: 4

US-10-751-235-1 (1-77) x US-11-128-061-473 (1-2858)
QY 6 GluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIle 25
Db 1410 GAGAAAGGGAAGCCGTCATATGAAGAAGAAATCCTGGAGCTTATAGCATGCGATCGATC 1469
QY 26 GlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThr---AspSerProValIle 44
Db 1470 ACTGGCACATCATTTGGAGTGAATGCGATTCCTCAACCAACCCAGAGGATCCCTTTGTG 1529
QY 45 GluAlaValTyrThrAlaLeuLys-----GluAlaGluLeuArgSerThrAsp 60
Db 1530 CAGAAGGCCAGGAAGATTTAAGCTTAATTTTGTATCCCTTATCTCTCCATAATA 1589
QY 61 LeuLeuProTyrTrpLys-----IleAspAlaLeu---CysLysIleValProArgGln 77
Db 1590 CTGTTTCCATTCCTTACCACATATATGACTGTTTAAGGTTCTCCATTTTCCAGACAA 1649

RESULT 14
US-10-240-708-27
; Sequence 27, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/240,708
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/EP01/039971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 473
; LENGTH: 2858
; TYPE: DNA
; ORGANISM: Cricetinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (129)..(191)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (366)..(400)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (481)..(501)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1020)..(1034)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-473
Alignment Scores:
Pred. No.: 266 Length: 2858
Score: 61.00 Matches: 25
Percent Similarity: 43.75% Conservative: 10
Best Local Similarity: 31.25% Mismatches: 37
Query Match: 15.52% Indels: 8
DB: 7 Gaps: 4

US-10-751-235-1 (1-77) x US-11-128-061-473 (1-2858)
QY 6 GluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIle 25
Db 1410 GAGAAAGGGAAGCCGTCATATGAAGAAGAAATCCTGGAGCTTATAGCATGCGATCGATC 1469
QY 26 GlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThr---AspSerProValIle 44
Db 1470 ACTGGCACATCATTTGGAGTGAATGCGATTCCTCAACCAACCCAGAGGATCCCTTTGTG 1529
QY 45 GluAlaValTyrThrAlaLeuLys-----GluAlaGluLeuArgSerThrAsp 60
Db 1530 CAGAAGGCCAGGAAGATTTAAGCTTAATTTTGTATCCCTTATCTCTCCATAATA 1589
QY 61 LeuLeuProTyrTrpLys-----IleAspAlaLeu---CysLysIleValProArgGln 77
Db 1590 CTGTTTCCATTCCTTACCACATATATGACTGTTTAAGGTTCTCCATTTTCCAGACAA 1649

; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 27
; LENGTH: 11131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-27
Alignment Scores:
Pred. No.: 1.85e+03 Length: 11131
Score: 61.00 Matches: 17
Percent Similarity: 49.06% Conservative: 9
Best Local Similarity: 32.08% Mismatches: 15
Query Match: 15.52% Indels: 12
DB: 6 Gaps: 2

US-10-751-235-1 (1-77) x US-10-240-708-27 (1-11131)
QY 19 GlnMetThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThr 38
Db 106 AGGATGTTTTTATAAATTTTAGGAAGGTATTATTTTATTTTAAATTTTAAAGGA----- 159
QY 39 ThrAspSerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSer 58
Db 160 -----ACGTATTATATAGAGTGTGTTATTTAAATATGATAGAGAACGCGAG----- 204
QY 59 ThrAspLeuLeuProTyrTrpLysIleAspAlaLeuCys 71
Db 205 -----TTTGGAAAGTTAGATTGCTTTGTGT 228

RESULT 15
US-10-793-626-1623
; Sequence 1623, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1623
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1623
Alignment Scores:
Pred. No.: 126 Length: 1494
Score: 60.50 Matches: 21
Percent Similarity: 38.71% Conservative: 15
Best Local Similarity: 22.58% Mismatches: 20
Query Match: 15.39% Indels: 37
DB: 6 Gaps: 4

US-10-751-235-1 (1-77) x US-10-793-626-1623 (1-1494)
QY 3 ProTyrAlaGluAspGlySer-----AlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 139 CCATTTCAAAATGATAGTAGCGAAGTACTATTATTAACATGCCATCAAAATCCATGAGCTTA 198
QY 21 ThrLeuAsp----- 23
```

```

Db      199 ANTCITGATGATCAAGAGTTTGGAAAGTGGTATCAAAAATCAGACGGAATTTAATTTT 258
Qy      24  |||||
Db      259 AGTAATCCTCAATATTTGCTAGATTGTATTTGGTCATTATATGAAGTCTTATTTAICT 318
Qy      29  LeuPheAsnTyrAsnPheAspSerLeuThrThrAppSerProValIleGluAlaValTyr 48
Db      319 TATTATAATGACCAATTTTGATTAATTTAACTATTATCAATGATAAAGTACACAGAAATTTT 378
Qy      49  ThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu 61
Db      379 ACA-----CAATCCGATGTTGATGACACAGATTTA 408

```

Search completed: January 15, 2006, 15:29:28
 Job time : 249 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 05:07:45 ; Search time 474 Seconds
(without alignments)
1082.661 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYADGSAVMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgm2.1/USPTO spool/US10751235/runat_12012006_120915_2508/app_query.fasta.1.263
-DB=N Geneseq -QFMT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CGN 1 1 727 @runat_12012006_120915_2508 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	3377	13 ADT15447	Adt15447 Plant cDN
2	346	88.0	2467	14 AEB16909	Aeb16909 Thale cre
3	346	88.0	2467	14 AEB16911	Aeb16911 Thale cre
4	346	88.0	4170	14 AEB16910	Aeb16910 Thale cre

ALIGNMENTS

RESULT 1

ADT15447

ID ADT15447 standard; cDNA; 3377 BP.

XX AC ADT15447;

XX DT 13-JAN-2005 (first entry)

XX DE Plant cDNA, Seq ID 773.

XX KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
XX OS Viridiplantae.
XX FN US2004216190-A1.
XX PD 28-OCT-2004.
XX PF 18-DEC-2003; 2003US-00739930.
XX PR 28-APR-2003; 2003US-00424599.
XX PP 28-APR-2003; 2003US-00425115.

Aeb16928 Maize cyt
Adx61765 Plant ful
Aeb16927 Tomato cy
Aeb16926 Wheat cyt
Aeb16925 Barley cy
Aeb16924 Rice cyto
Aeb16929 Sunflower
Aeb16950 Green alg
Aeb16948 Wheat cyt
Aeb16946 Barley cy
Aeb16945 Rice cyto
Ado81358 Plant ful
Aeb16957 Soybean c
Aax60780 Soybean c
Aeb16943 Thale cre
Aeb16942 Thale cre
Aaag8063 Pinus rad
Add41813 Cinnamat
Aeb16947 Soybean c
Aeb16955 Thale cre
Aeb16949 Tomato cy
Aeb16956 Pea CYP97
Aeb16944 Rice cyto
Aeb16959 Diatom cy
Ady27372 Bacillus
Ady27460 Bacillus
Ady27356 Bacillus
Ady27458 Bacillus
Abk74941 Bacillus
Aag02937 Heme dona
Ady27354 Bacillus
Aag02940 Hybrid pr
Ady27456 Bacillus
Aaa46648 Nucleotid
Aad02365 Bacillus
Aaf54832 Nucleotid
Aaf26288 B. megate
Aaf26324 B. megate
Acc48000 B. megata
Abz70676 Bacillus
Aal54548 DNA of P4

PA (KOVA/) KOVALIC D K.
XX Kovalic DK;
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
XX genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.
XX
XX Claim 1; SEQ ID NO 773; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
XX polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
XX 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
XX (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
XX Arabidopsis, wheat and rape but the specification does not indicate which
XX sequences is derived from which organism. Also included is a method of
XX producing a plant having an improved property, comprising transforming a
XX plant with a recombinant DNA construct comprising a promoter region
XX functional in a plant cell operably joined to a polynucleotide encoding a
XX polypeptide associated with the property, and growing the transformed
XX plant. The property is selected from improving plant cold tolerance, for
XX manipulating growth rate in plant cells by modification of the cell cycle
XX pathway, for improving plant drought tolerance, for providing increased
XX resistance to plant disease, for galactomannan production, for production
XX of plant growth regulators, for improving plant heat tolerance, for
XX improving plant tolerance to herbicides, for increasing the rate of
XX homologous recombination in plants, for lignin production, for improving
XX plant tolerance to extreme osmotic conditions, for improving plant
XX tolerance to pathogens or pests, for yield improvement by modification of
XX photosynthesis, for modifying seed oil yield and/or content, for
XX modifying seed protein yield and/or content, for yield improvement by
XX modification of carbohydrate, nitrogen or phosphorus use and/or uptake
XX and for yield improvement by providing improved plant growth and
XX development under at least one stress condition. The polynucleotide may
XX also encode a plant transcription factor. The methods and compositions of
XX the present invention are useful in the field of biochemistry and
XX genetics, in particular for producing transgenic plants with improved
XX biological characteristics such as increased yield, improved nitrogen
XX flow, increasing plant tolerance to cold or heat, improving plant
XX tolerance to extreme osmotic and drought conditions, and improving plant
XX tolerance to plant pests or pathogens. They can also be used in physical
XX arrays of molecules, plant breeding markers, computer-based storage and
XX analysis systems. The present sequence is one of the 5544 plant CDNA
XX sequences of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
XX Sequence 3377 BP; 967 A; 729 C; 736 G; 945 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,3e-49 Length: 3377
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-751-235-1 (1-77) x APT15447 (1-3377)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 676 TTGCACCTTATGCAGACGCGAAGTGTGTGAATATGAGCGAAGTCTCTCAGATG 735
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 736 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAAATTCGATTCCTTGACTACTGAT 795
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 796 AGTCCTGTCTATTGAAGCTGTTTACATGCTCTTTAAGAGCTGAGTCTGCTTCTACTGAT 855

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 856 CTTCTGCCATATTGGAAGATCGATTCATTGTGTAAATATGTCCTCCGACAG 906
RESULT 2
ID AEB16909 standard; cDNA; 2467 BP.
XX AEB16909;
XX
XX 08-SEP-2005 (first entry)
XX
XX Thale cress LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 5.
XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
XX cytochrome P450 97C1; gene; ss; chromosome 3.
XX Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
XX CDS 1..2467
XX /*tag= a
XX /product= "Thale cress LUT1 protein"
XX /transl_except= (pos:496..650, aa:Thr-Ala)
XX /transl_except= (pos:942..1037, aa:Lys-Ile)
XX /transl_except= (pos:1242..1326, aa:Val-Ser)
XX /transl_except= (pos:1420..1523, aa:Lys-Asn)
XX /transl_except= (pos:1662..1841, aa:Val-Leu)
XX /transl_except= (pos:1839..2005, aa:Val-Trip)
XX /transl_except= (pos:2081..2183, aa:Lys-Phe)
XX /transl_except= (pos:2340..2440, aa:Asn-Gly)
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIM/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16908.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 5; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in altering the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is thale cress LUT1 (cytochrome P450
XX monooxygenase (CYP97C1); AEB16908 gene) cDNA. Note: The current sequence
XX is that of thale cress LUT1 cDNA which is located on chromosome 3.
XX
XX Sequence 2467 BP; 657 A; 491 C; 521 G; 798 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.79e-42 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservatives: 0

Best Local Similarity: 71.03% Mismatches: 1
 Query Match: 88.04% Indels: 31
 DB: 14 Gaps: 1

US-10-751-235-1 (1-77) x AEB16909 (1-2467)

CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is thale cress LUT1 [cytochrome P450
 CC monooxygenase (CYP97C1)] mutant cDNA.

XX SQ Sequence 2467 BP; 658 A; 491 C; 520 G; 798 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 7,798-42 Length: 2467
 Score: 346.00 Matches: 76
 Percent Similarity: 71.03% Conservative: 0
 Best Local Similarity: 71.03% Mismatches: 1
 Query Match: 88.04% Indels: 31
 DB: 14 Gaps: 1

US-10-751-235-1 (1-77) x AEB16911 (1-2467)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 DB 747 TTGCAGCCTTATGCAAGACGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATG 806

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 DB 807 ACACCTGATGTCATTTGGTGTCTCTTTTAACTACATTTTCGATCTTTGACTACTGAT 866

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 867 AGTCCTGTCTTGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTTCGTTCTACTGAT 926

QY 61 LeuLeuProTyrTropLysIle----- 67
 DB 927 CTTCTGCCCATATTTGGAAGGC-AAGTTTCTGTGTTTTTTTCTGTGTTTGTGTTGATTTG 985

QY 68 -----AspAlaLeu 70
 DB 986 GAACAATTGGATCTTGTTAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGCAATG 1045

QY 71 CysLysIleValProArgGln 77
 DB 1046 TGTAAGATAGTCCCGAGACAG 1066

RESULT 4
 AEB16910
 ID AEB16910 standard; cDNA; 2467 BP.
 AC AEB16910;
 XX
 XX 08-SEP-2005 (first entry)
 DT
 XX
 DE Thale cress LUT1 (cytochrome P450 97C1) DNA, SEQ ID NO: 6.
 XX
 KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
 KW cytochrome P450 97C1; gene; ds; chromosome 3.
 XX
 OS Arabidopsis thaliana.
 XX
 XX US2005150002-A1.
 XX
 XX 07-JUL-2005.
 PD
 XX
 PF 02-JAN-2004; 2004US-00751235.
 XX
 PR 02-JAN-2004; 2004US-00751235.
 XX
 PA (DELL/) DELLAPENNA D.
 PA (TIAN/) TIAN L.
 PA (KIMJ/) KIM J.
 XX
 PI Dellapenna D, Tian L, Kim J;
 XX
 DR WPI; 2005-487984/49.
 XX
 PT New expression vector comprising a nucleic acid sequence encoding a

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 DB 747 TTGCAGCCTTATGCAAGACGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATG 806

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 DB 807 ACACCTGATGTCATTTGGTGTCTCTTTTAACTACATTTTCGATCTTTGACTACTGAT 866

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 867 AGTCCTGTCTTGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTTCGTTCTACTGAT 926

QY 61 LeuLeuProTyrTropLysIle----- 67
 DB 927 CTTCTGCCCATATTTGGAAGGC-AAGTTTCTGTGTTTTTTTCTGTGTTTGTGTTGATTTG 985

QY 68 -----AspAlaLeu 70
 DB 986 GAACAATTGGATCTTGTTAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGCAATG 1045

QY 71 CysLysIleValProArgGln 77
 DB 1046 TGTAAGATAGTCCCGAGACAG 1066

RESULT 3
 AEB16911
 ID AEB16911 standard; cDNA; 2467 BP.
 AC AEB16911;
 XX
 XX 08-SEP-2005 (first entry)
 DT
 XX
 DE Thale cress mutant LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 7.
 XX
 KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
 KW cytochrome P450 97C1; ss; mutant.
 XX
 OS Arabidopsis thaliana.
 OS Synthetic.
 XX
 XX US2005150002-A1.
 XX
 XX 07-JUL-2005.
 PD
 XX
 PF 02-JAN-2004; 2004US-00751235.
 XX
 PR 02-JAN-2004; 2004US-00751235.
 XX
 PA (DELL/) DELLAPENNA D.
 PA (TIAN/) TIAN L.
 PA (KIMJ/) KIM J.
 XX
 PI Dellapenna D, Tian L, Kim J;
 XX
 DR WPI; 2005-487984/49.
 XX
 PT New expression vector comprising a nucleic acid sequence encoding a
 PT polypeptide having monooxygenase P450 activity, useful in altering the
 PT carotenoid production in a plant for enhancing production of specific
 PT carotenoid compounds.
 XX
 PS Claim 10; SEQ ID NO 7; 135pp; English.
 XX
 CC The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The

PT polypeptide having monooxygenase P450 activity, useful in altering the carotenoid production in a plant for enhancing production of specific carotenoid compounds.

XX Claim 10; SEQ ID NO 6; 135pp; English.

XX The present invention relates to genes, proteins and methods comprising carotenoid monooxygenases in the cytochrome P450 family. The invention also relates to altering carotenoid ratios in plants and microorganisms using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The invention is useful in altering the carotenoid production in a plant for enhancing production of specific carotenoid compounds that are potent antioxidants. The present sequence is thale cress LUT1 [cytochrome P450 monooxygenase (CYP97C1); At3g53130 gene] genomic DNA. Note: The current sequence is that of thale cress LUT1 DNA which is located on chromosome 3.

XX Sequence 4170 BP; 1222 A; 862 C; 812 G; 1274 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.63e-41 Length: 4170
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: 14 Gaps: 1

US-10-751-235-1 (1-77) x AEB16910 (1-4170)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
1756 TTGCAGCCTTATGCGAAGACGGAAGTGTGTGAATATGGAAGCGAAGTCTCTCAGATG 1815

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
1816 ACATTCGATGTCATTGGGTGTCCTTTTAACTACATTCGATTCCTTGACTACTGAT 1875

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
1876 AGTCCTGTGTCATTGAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 1935

QY 61 LeuLeuProTyrTriPlysile----- 67
1936 CTTCTGCCATATGGAAGGC-AAGTTTCTGTGTTTTTTCTGTGGTTTGTGATGTTGTG 1994

QY 68 -----AspAlaLeu 70
1995 GAACAATTGATCTCTGTTAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGCAATTG 2054

QY 71 CysLysIleValProArgGln 77
2055 TGTAAGATAGTCCCGAGACAG 2075

RESULT 5
AEB16928
ID AEB16928 standard; cDNA; 531 BP.
XX
AC AEB16928;
XX
DT 08-SEP-2005 (first entry)
XX
DE Maize cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 26.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Zea mays.
XX
FH key Location/Qualifiers
FT CDS 1..531
FT /tag= a
FT /product= "Maize cytochrome P450 97C (CYP97C) protein"
FT /partial

FT
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-007511235.
XX
PR 02-JAN-2004; 2004US-007511235.
XX
PA (DELL/) DELLAPENNA D.
(TIAN/) TIAN L.
(KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
P-PSDB; AEB16922.
DR
DR GENBANK; BE552887.
XX
PS New expression vector comprising a nucleic acid sequence encoding a polypeptide having monooxygenase P450 activity, useful in altering the carotenoid production in a plant for enhancing production of specific carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 26; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising carotenoid monooxygenases in the cytochrome P450 family. The invention also relates to altering carotenoid ratios in plants and microorganisms using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The invention is useful in altering the carotenoid production in a plant for enhancing production of specific carotenoid compounds that are potent antioxidants. The present sequence is maize cytochrome P450 monooxygenase (CYP97C) cDNA.
XX
SQ Sequence 531 BP; 158 A; 91 C; 130 G; 152 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.59e-42 Length: 531
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 6
Query Match: 87.28% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16928 (1-531)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
1 CTTGAGCCATATGCTTTGAGTGGGAACTGTCATATGGAAGCGAGGTTTTCTCAGTTG 60

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
61 ACATTGGATGATGTTTATCATCTGTTCAACTACAATTTGATTCCTCCACACAGAT 120

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
121 AGTCCTGTGCTGTTCTTTATCTGCACTCAAGAAGCAGAGTTCTGTTCTACAGAT 180

QY 61 LeuLeuProTyrTriPlysileAspAlaLeuCysLysIleValProArgGln 77
181 CTTTGGCCATCTGGAGGTTGTTTCTTGTGCAAGATAATCCCAAGACAG 231

RESULT 6
ADX61765
ID ADX61765 standard; cDNA; 1804 BP.
XX
AC ADX61765;
XX
DT 21-APR-2005 (first entry)
XX
DE plant full length insert polynucleotide seqid 32608.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX Unidentified.
 OS US2004034888-A1.
 XX PD 19-FEB-2004.
 XX PF 28-APR-2003; 2003US-00425114.
 XX PR 06-MAY-1999; 99US-00304517.
 XX PR 05-NOV-2001; 2001US-00985678.
 XX (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI; 2004-180133/17.
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX Claim 1; SEQ ID NO 32608; 15pp; English.
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX Sequence 1804 BP; 464 A; 392 C; 472 G; 476 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 1,44e-41 Length: 1804
 Score: 343.00 Matches: 65
 Percent Similarity: 92.21% Conservative: 6
 Best Local Similarity: 84.42% Mismatches: 6
 Query Match: 87.28% Indels: 0
 DB: 13 Gaps: 0

US-10-751-235-1 (1-77) x ADX61765 (1-1804)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 |||||
 DB 494 CTTGAGCCATATGCTTTGTAGTGGGAACCTGTCAATATGGAGCGAGGTTTCTCAGTTG 553
 |||||

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 |||||
 DB 554 ACATTGGATGTGATTGTTTATCATTTGTTCACTACAAATTTTGTATTCCCTCCACACAGAT 613
 |||||
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 |||||
 DB 614 AGTCCTGTGATGATGCTGTTTATATCTGCACCTCAAGAAGCAGAGCTTCGTTCTACAGAT 673
 |||||
 QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCyLSysIleValProArgGln 77
 |||||
 DB 674 CTTTGGCATACCTGGAAGGTGTTTCTTGTGCAAGTAAATCCCAAGACAG 724
 |||||

RESULT 7
 AEB16927
 ID AEB16927 standard; cDNA; 839 BP.
 XX AC AEB16927;
 XX DT 08-SEP-2005 (first entry)
 XX DE Tomato cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 25.
 XX KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
 KW cytochrome P450 97C; Gene; ss.
 XX OS Lycopersicon esculentum.
 XX FH Key Location/Qualifiers
 CDS 1..837
 FT /*tag= a
 FT /product= "Tomato cytochrome P450 97C (CYP97C) protein"
 FT /partial
 FT /note= "No start and stop codons"
 XX PN US2005150002-A1.
 XX PD 07-JUL-2005.
 XX PF 02-JAN-2004; 2004US-00751235.
 XX PR 02-JAN-2004; 2004US-00751235.
 XX PA (DELL/) DELLAPENNA D.
 PA (TIAN/) TIAN L.
 PA (KIM/) KIM J.
 XX PI Dellapenna D, Tian L, Kim J;
 XX WPI; 2005-487984/49.
 DR P-PSDB; AEB16921.
 DR GENBANK; BG643819.
 XX PT New expression vector comprising a nucleic acid sequence encoding a
 PT polypeptide having monooxygenase P450 activity, useful in altering the
 PT carotenoid production in a plant for enhancing production of specific
 PT carotenoid compounds.
 XX PS Claim 10; SEQ ID NO 25; 135pp; English.
 XX CC The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using lutein epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
 CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is tomato cytochrome P450
 CC monooxygenase (CYP97C) cDNA.
 XX SQ Sequence 839 BP; 251 A; 160 C; 190 G; 238 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.04e-41 Length: 839

Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16927 (1-839)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 31 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGGAGGCAAAAGTTTCTCAACTA 90

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 91 ACACCTGATGTTATGTCCTTGCACCTCTCAATTAATTTTGGATTCCTTACTACTGAC 150

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 151 AGTCAGTTATGATGAGCTTTACACTGCATTAAGAAGCAGCAACTCGTTCAACTGAT 210

QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
Db 211 TTGTTGCCATATTGGCAGATCAAAGCTTTATGTAAGTTTCATCCACGACAA 261

RESULT 8
AEB16926
ID AEB16926 standard; cDNA; 1086 BP.
XX
AC AEB16926;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 24.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 1..1086
FT /tag= a
FT /product= "wheat cytochrome P450 97C (CYP97C) protein"
FT /partial
FT /note= "No stop codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16920.
DR GENBANK; CA497665, BG906289, CA742365, CA742792.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 24; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention

CC also relates to altering carotenoid ratios in plants and microorganisms
CC using IUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC (CYP97C) cDNA.

XX Sequence 1086 BP; 243 A; 298 C; 295 G; 250 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.17e-40 Length: 1086
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16926 (1-1086)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 589 CTCGAGACTTATGCTTTCAGTGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAATG 648

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 649 ACATTAGATGATGTTGTTTATCCTTCAACTACAACTTTGATTCCTCCATCAGAT 708

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 709 AGTCCTGTTATGATGCTTTTACACTGCACCTCAAGAAGCTGAGGCTGTTCTACAGAT 768

QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
Db 769 CTTTACCATCTGCAGATCGATTTGCTGTGCAAGATTTCTCTAGACAG 819

RESULT 9

AEB16925
ID AEB16925 standard; cDNA; 1638 BP.
XX
AC AEB16925;
XX
DT 08-SEP-2005 (first entry)
XX
DE Barley cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 23.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Hordeum vulgare; subsp. vulgare.
OS Hordeum vulgare; subsp. spontaneum.
XX
FH Key Location/Qualifiers
FT CDS 1..1638
FT /tag= a
FT /product= "Barley cytochrome P450 97C (CYP97C) protein"
FT /transl_except= (pos:1585..1587, aa:Xaa)
FT /note= "Xaa may be any naturally occurring amino acid"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.

```

DR P-PSDB; AEB16919.
DR GENBANK; BM816653, BU987393, CA023004.
DR DDBJ; AV835803.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 23; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX
SQ Sequence 1638 BP; 406 A; 403 C; 427 G; 401 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2,08e-40 Length: 1638
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16925 (1-1638)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 574 CTGAGACATATGCTTTGAGCGGTGAACCTGTTAATATGGAAGCGAGATTTCTCAATG 633
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrApp 40
Db 634 ACACATAGATGTGATGGTTGTTGTTGTTCACTCAACTTGAATTCCTCATCAGAT 693
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrApp 60
Db 694 AGTCCTGTTATGATGCTGTTTACACCGCACTGGAAGAGCAGAGGCTGTTCTACAGAT 753
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 754 CTTTACCATACTGGCAGATTCGATTCGTCGTCGACAGATTGTTCTCAGACAG 804

RESULT 10
AEB16924
ID AEB16924 standard; cDNA; 1686 BP.
XX
AC AEB16924;
XX
DT 08-SEP-2005 (first entry)
XX
DE Rice cytochrome P450 97C2 (CYP97C2) cDNA, SEQ ID NO: 22.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C2;
KW cytochrome P450 97C2; gene; ss.
XX
OS Oryza sativa; japonica cultivar-group.
XX
FH Key Location/Qualifiers
FT CDS
FT 1..1686
FT /tag=a
FT /product="Rice cytochrome P450 97C2 (CYP97C2) protein"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.

XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16918.
XX DDBJ; AK065689.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 22; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC (CYP97C2) cDNA.
XX
XX Sequence 1686 BP; 393 A; 415 C; 459 G; 419 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,17e-40 Length: 1686
Score: 335.00 Matches: 66
Percent Similarity: 90.91% Conservative: 4
Best Local Similarity: 85.71% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16924 (1-1686)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTTAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrApp 40
Db 682 ACTTTAGATGTGATGGTTGTTGTTGTTCAATTACAAATTTTGATTCCTCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrApp 60
Db 742 AGCCCTGTTATGATGCTGTTTACACTGCACCTCAAGGAAGCAAGAACTTCGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAGATTCGATTCGTCGTCGACAGATTGTTCTTAGACAA 852

RESULT 11
AEB16929
ID AEB16929 standard; cDNA; 624 BP.
XX
AC AEB16929;
XX
DT 08-SEP-2005 (first entry)
XX
DE Sunflower cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 27.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
XX Helianthus annuus.
XX
XX Key Location/Qualifiers

```

```
FT CDS 1..624
FT /*tag= a
FT /product= "Sunflower cytochrome P450 97C (CYP97C)
FT /protein=
FT /partial
FT /note= "No start and stop codons"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16923.
XX GENBANK, BQ971938.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 27; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in affecting the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is sunflower cytochrome P450
XX monooxygenase (CYP97C) cDNA.
XX
XX Sequence 624 BP; 186 A; 110 C; 144 G; 184 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.48e-38 Length: 624
XX Score: 319.00 Matches: 60
XX Percent Similarity: 89.61% Conservative: 9
XX Best Local Similarity: 77.92% Mismatches: 8
XX Query Match: 81.17% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-751-235-1 (1-77) x AEB16929 (1-624)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 253 CTAAGATCATACCGACGACGTGCTGTTAATGAGCAGACAGTTTCGCGATTA 312
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 313 ACCCTTGATGTTATCGTCTAGCGGTATTAACTACAAATTTTCACTTACGCGCAT 372
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 373 AGTCCTGTAATCGAATCTGTTTATACCGCACTAAAGAGAGCTGAAGCCGCTTCACTGAT 432
QY 61 LeuLeuProTyrTpolysIleAspAlaLeuCysIysIleValProArgGln 77
DB 433 CTTTGGCATATGGAGAGATAGTGGCTTATGTAAGATTATACAGACAA 483
XX
XX RESULT 12
XX AEB16950
XX ID AEB16950 standard; cDNA; 1101 BP.
XX
```

```
AC AEB16950;
XX
XX 08-SEP-2005 (first entry)
XX
XX Green algae cytochrome P450 97A (CYP97A) like cDNA, SEQ ID NO: 48.
XX
XX DE Green algae cytochrome P450 97A (CYP97A) like cDNA, SEQ ID NO: 48.
XX
XX KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ss.
XX
XX OS Chlamydomonas reinhardtii.
XX
XX Key Location/Qualifiers
XX CDS 1..1101
XX FT /*tag= a
XX FT /product= "Green algae cytochrome P450 97A (CYP97A) like
XX FT protein"
XX FT /partial
XX FT /note= "No start and stop codons"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16941.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 48; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in affecting the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is green algae cytochrome P450
XX monooxygenase (CYP97A) like cDNA.
XX
XX Sequence 1101 BP; 217 A; 353 C; 342 G; 189 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.35e-21 Length: 1101
XX Score: 210.00 Matches: 41
XX Percent Similarity: 72.73% Conservative: 15
XX Best Local Similarity: 53.25% Mismatches: 21
XX Query Match: 53.44% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-751-235-1 (1-77) x AEB16950 (1-1101)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 100 CTAGACAAGTATGCGCGCTCAGGCACCGCTGGACATGGAACACTTCTTCACCGCGCTG 159
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 160 GGTCTGACATCATCGCGAGCCGTTGTTCACTACACTTCCGCTCGCTGCGCGCAGAC 219
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
```

```
Db 220 GACCCGCTATCAGCCGCTGACACCTTCTGCGGAGCGAGACCGCTCCACAGCG 279
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 280 CCATCGCCTACTGGAACATTCGCGCATCCAGTTTGTGTGTCGCGGCAG 330

RESULT 13
AEB16948
ID AEB16948 standard; cDNA; 1031 BP.
XX
AC AEB16948;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97A (CYP97A) cDNA, SEQ ID NO: 46.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 1..1031
FT /tag= a
FT /product= "Wheat cytochrome P450 97A (CYP97A) protein"
FT /transl_except= (pos:592..595, aa:Gln)
FT /partial
FT /note= "No start codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16939.
DR DDBJ; BJ234910.
DR GENEANK; CA736787, CA736801.
DR DDBJ; BJ238659, BJ233019.
DR GENEANK; CD882035.
XX
New expression vector comprising a nucleic acid sequence encoding a
polyptide having monooxygenase P450 activity, useful in altering the
carotenoid production in a plant for enhancing production of specific
carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 46; 135pp; English.
XX
The present invention relates to genes, proteins and methods comprising
carotenoid monooxygenases in the cytochrome P450 family. The invention
also relates to altering carotenoid ratios in plants and microorganisms
using luti epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
invention is useful in altering the carotenoid production in a plant for
enhancing production of specific carotenoid compounds that are potent
antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
(CYP97A) cDNA.
XX
SQ Sequence 1031 BP; 253 A; 237 C; 284 G; 257 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,52e-21 Length: 1031
Score: 207.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
```

```
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.67% Indels: 0
DB: 14 Gaps: 0
US-10-751-235-1 (1-77) x AEB16948 (1-1031)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 364 CTGGACAAGCGGCATCCGATGGGAGGATGTGGAGATGGAAATCTCTCTCTCGACTA 423
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 424 ACGCTGGATGTCATCGGAAGCAGTGTCAATATGATTTTGATTCATTATCTTACGAT 483
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 484 AATGGAATAGTTGAGGCTGTGTATGTAACATTACGGGAAGCGGAAATGCGGAGCATCT 543
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 544 CCTATTCCAACCTGGGAATAATCCCATATGGAAGACATCTCCCTCGGCAG 594

RESULT 14
AEB16946
ID AEB16946 standard; cDNA; 1527 BP.
XX
AC AEB16946;
XX
DT 08-SEP-2005 (first entry)
XX
DE Barley cytochrome P450 97A (CYP97A) cDNA, SEQ ID NO: 44.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ss.
XX
OS Hordeum vulgare; subsp. vulgare.
XX
FH Key Location/Qualifiers
FT CDS 1..1527
FT /tag= a
FT /product= "Barley cytochrome P450 97A (CYP97A) protein"
FT /partial
FT /note= "No start codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16937.
DR DDBJ; AV939715, AV941342, AV939552, AV939556.
DR GENEANK; CA004011.
DR DDBJ; BJ480615, BJ485000, BJ448041, BJ455787, AV910152, AV938407.
DR EMBL; AJ477620, AJ477618, AJ477619, AJ832622.
XX
New expression vector comprising a nucleic acid sequence encoding a
polyptide having monooxygenase P450 activity, useful in altering the
carotenoid production in a plant for enhancing production of specific
carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 44; 135pp; English.
XX
```

CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97A) cDNA.

XX
XX
SQ Sequence 1527 BP; 399 A; 343 C; 395 G; 390 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.11e-21 Length: 1527
Score: 207.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.67% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16946 (1-1527)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 361 CTCGACAAAGCGCTCTCGACGCGGAGGATGGAGATGGAACTCTCTTCTCCGACTA 420

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 421 AGCTGGATGTCATCGGGAAGCGGTGTTCAATTATGATTTTGAATTCATTATCTTACGAT 480

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 481 AATGGAATAGTTGAGCTGTGTATGTAACACTGCGGGAAGCAAGAAATGCGGAGTACATCT 540

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 541 CCTATTCCAACTGGGAATACCCATATGGAAAGACATCTCCCTCGGCAG 591

RESULT 15
AEB16945
ID AEB16945 standard; DNA; 1899 BP.

XX
AC AEB16945;
DT
XX 08-SEP-2005 (first entry)

DE Rice cytochrome P450 97A (CYP97A) DNA, SEQ ID NO: 43.

XX
XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ds.
XX
XX Oryza sativa; japonica cultivar-group.

XX
FH Key Location/Qualifiers
FT CDS 1..1899
FT /tag= a
FT /product= "Rice cytochrome P450 97A (CYP97A) protein"

XX
FN US2005150002-A1.
XX
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX
XX P-PSDB; AEB16936.

DR DDBJ; AP004028.

XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.

XX
XX Claim 10; SEQ ID NO 43; 135pp; English.

XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC (CYP97A) DNA.

XX
XX Sequence 1899 BP; 454 A; 470 C; 504 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.18e-20 Length: 1899
Score: 206.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.42% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16945 (1-1899)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 733 TTGGACAAGGCGAGCAACGATGGGAGGATGGAGATGGAACTCTTCTCTCGACTA 792

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 793 ACACCTGGATGTCATTGGGAAGCGACTTCAATTATGATTTTCGACTCATTGCTTACGAT 852

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 853 AATGGAATAGTTGAGGCGAGTGTATGTGACACTGCGGGAAGCAAGAAATGCGGAGCACTTCT 912

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 913 CCTATACCAACTTGGGAATACCCATATGGAAAGATATTTCCTCCGCGGCAG 963

Search completed: January 15, 2006, 13:20:16
Job time : 477 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 12:59:30 ; Search time 171 Seconds
(without alignments)
800.423 Million cell updates/sec

Title: US-10-751-235-1
Perfect score: 393
Sequence: 1 LQVABDSAVNMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-DB=Issued Patents NA -QWNT=fastap -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CGN 1 1 193 @runat 12012006 120916 2538 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgm2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/1/ina/H COMB.seq.*
6: /cgm2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgm2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgm2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgm2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	200	50.9	1831 3	US-08-948-564-15 Sequence 15, Appl
2	196	49.9	404 3	US-09-615-192A-156 Sequence 156, App
3	196	49.9	404 3	US-09-169-789-156 Sequence 156, App
4	76	19.3	3150 3	US-10-018-730A-3 Sequence 3, Appli
5	74	18.8	1633 3	US-09-583-447A-7 Sequence 7, Appli
6	71.5	18.2	1349 3	US-09-583-447A-5 Sequence 5, Appli
7	71.5	18.2	1515 3	US-09-583-447A-3 Sequence 3, Appli
8	71.5	18.2	1659 3	US-09-583-447A-1 Sequence 1, Appli
9	71.5	18.2	1973 3	US-09-583-447A-11 Sequence 11, Appl

10	71	18.1	2352 2	US-08-802-141-3 Sequence 3, Appli
11	71	18.1	6176 3	US-08-911-321-6 Sequence 6, Appli
12	69.5	17.7	432 3	US-09-248-796A-4077 Sequence 4077, Ap
13	69	17.6	2217 3	US-09-712-363-92 Sequence 92, Appl
14	69	17.6	2363 2	US-07-603-133B-19 Sequence 19, Appl
15	69	17.6	2368 2	US-08-653-740-2 Sequence 2, Appli
16	69	17.6	2368 2	US-09-073-594-2 Sequence 2, Appli
17	69	17.6	2368 3	US-09-275-925-2 Sequence 2, Appli
c 18	69	17.6	4403765 3	US-09-103-840A-2 Sequence 2, Appli
c 19	69	17.6	441529 3	US-09-103-840A-1 Sequence 1, Appli
20	65.5	16.7	2338 3	US-09-336-115C-21 Sequence 21, Appl
21	65.5	16.7	2429 3	US-09-336-115C-11 Sequence 11, Appl
22	65.5	16.7	73788 3	US-09-949-016-12358 Sequence 12358, A
23	65	16.5	629 2	US-08-454-557C-43 Sequence 43, Appl
24	65	16.5	629 2	US-08-340-426D-43 Sequence 43, Appl
25	65	16.5	629 2	US-08-450-673C-43 Sequence 43, Appl
26	65	16.5	629 6	PCT-US95-17111A-43 Sequence 43, Appl
27	65	16.5	1431 3	US-09-268-992-6 Sequence 6, Appli
28	65	16.5	1431 3	US-09-657-474-6 Sequence 6, Appli
29	65	16.5	1485 3	US-09-268-992-5 Sequence 5, Appli
30	65	16.5	1485 3	US-09-657-474-5 Sequence 5, Appli
31	65	16.5	1957 3	US-09-268-992-3 Sequence 3, Appli
32	65	16.5	1957 3	US-09-657-474-3 Sequence 3, Appli
33	65	16.5	2055 3	US-09-268-992-1 Sequence 1, Appli
34	65	16.5	2055 3	US-09-657-474-1 Sequence 1, Appli
35	64.5	16.4	601 3	US-09-949-016-79444 Sequence 79444, A
36	64.5	16.4	784019 3	US-09-949-016-14033 Sequence 14033, A
37	64.5	16.4	828152 3	US-09-949-016-12777 Sequence 12777, A
38	64	16.3	804 3	US-09-710-279-1611 Sequence 1611, Ap
39	64	16.3	861 3	US-09-134-001C-2487 Sequence 2487, Ap
40	64	16.3	1707 3	US-09-023-655-1060 Sequence 1060, Ap
41	64	16.3	1707 3	US-09-949-016-121 Sequence 121, App
42	64	16.3	1707 3	US-09-949-016-1220 Sequence 1220, Ap
43	64	16.3	2918 3	US-09-710-279-3932 Sequence 3932, Ap
44	64	16.3	3660 3	US-09-710-279-4116 Sequence 4116, Ap
45	63.5	16.2	1158 3	US-09-543-681A-3780 Sequence 3780, Ap

ALIGNMENTS

RESULT 1

US-08-948-564-15
; Sequence 15, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400


```

; APPLICANT: Luet, Wong
; APPLICANT: Jonathan, Jones
; TITLE OF INVENTION: PROCESS FOR OXIDISING AROMATIC COMPOUNDS
; FILE REFERENCE: P02353US1 / 10112404 / N.76277B
; CURRENT APPLICATION NUMBER: US/10/018,730A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: GB 9914373.7
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/GB00/02379
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3150
; TYPE: DNA
; ORGANISM: Bacillus megaterium
US-10-018-730A-3

```

Alignment Scores:	
Pred. No.:	0.432
Score:	76.00
Percent Similarity:	56.52%
Best Local Similarity:	39.13%
Query Match:	19.34%
DB:	3
Length:	3150
Matches:	18
Conservative:	8
Mismatches:	18
Indels:	2
Gaps:	1

US-10-751-235-1 (1-77) x US-10-018-730A-3 (1-3150)

Qy	11	ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspVallIleGlyLeuSerLeuPhe	30
		::: :::	
Db	418	ATTGAAGTACCGGAAGACATGACACGGTTTAACGCTTGATACAATTGTCTTTTCGGCTTT	477
Qy	31	AsnTyrAsnPheAaspSerLeuThrThraspSer-----ProVallIleGluAlaValTyr	48
		:::	
Db	478	AACATATCGCTTTAACAGCTTTTACCGAGATCAGCGCTCATCCCAATTATTATACAAGTATGGTC	537
Qy	49	ThrAlaLeuLysGluAla	54
Db	538	CGTGACTGGATGAAGCA	555

RESULT 5

US-09-583-447A-7
; Sequence 7, Application US/09583447A

Patent No. 6645745
GENERAL INFORMATION:

; APPLICANT: WOJNOWSKI, Leszek
 ; APPLICANT: GELLNER, Klaus

;
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NE

```

; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3A4
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 7

Qy	1	LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet	20
Db	478	CTGAGCAGGAGGAGCAGAAACAGCAAGTCCATCACTTGAAGAATTTCTTTGGGGCCTAC	537
Qy	21	ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThr	39
Db	538	ACCATGATGTAATCACCTGGCACATATTCTGGAGTGAACCTTGGATCTCTCAACAATCCA	597
Qy	40	AspSerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThr	59
Db	598	CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA---TTGGATTTTTTGGATCCC	654
Qy	60	AspLeuLeuProTyrTrpLysIleAspAlaLeuCys	71
Db	655	TTTATTACTTAAATATACAGAGTCTCGCTGTGGTTC	690

RESULT 6

US-09-583-447A-5

; Sequence 5, Applicant
 ; Patent No. 6645745

;
; GENERAL INFORMATION:
;
; APPLICANT: WOJNOWSKI, Leszek

```

; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina

```

;; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYP3A GENE FAMILY: CYP3AX

FILE REFERENCE: 310115.401

```

;
; CURRENT FILING DATE: 2000-05-30
; CURRENT APPLICATION NUMBER: US/09/583,447A
;
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
;

```

SEQ ID	DESCRIPTION
1	...
2	...
3	...
4	...
5	...
6	...
7	...
8	...
9	...
10	...
11	...
12	...
13	...
14	...
15	...
16	...
17	...
18	...
19	...
20	...
21	...
22	...
23	...
24	...
25	...
26	...
27	...
28	...
29	...
30	...
31	...
32	...
33	...
34	...
35	...
36	...
37	...
38	...
39	...
40	...
41	...
42	...
43	...
44	...
45	...
46	...
47	...
48	...
49	...
50	...
51	...
52	...
53	...
54	...
55	...
56	...
57	...
58	...
59	...
60	...
61	...
62	...
63	...
64	...
65	...
66	...
67	...
68	...
69	...
70	...
71	...
72	...
73	...
74	...
75	...
76	...
77	...
78	...
79	...
80	...
81	...
82	...
83	...
84	...
85	...
86	...
87	...
88	...
89	...
90	...
91	...
92	...
93	...
94	...
95	...
96	...
97	...
98	...
99	...
100	...

LENGTH: 1349
TYPE: DNA

```

;
; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME / CDS

```

```

; NAME/KEY: CDS
; LOCATION: (1)..(1260)
;
; 00 503 4472 5

```

```

; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1515)
US-09-583-447A-3
Alignment Scores:
Pred. No.: 0.751 Length: 1515
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
DB: 3 Gaps: 1

US-10-751-235-1 (1-77) x US-09-583-447A-3 (1-1515)
Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 478 CTGAGCGAGGAGCAGACAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 537
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThr--- 39
Db 538 ACCATGATGTAATCACTGCGCACATTTATTTGGAGTGAACCTTGGATTCTCTCAACAATCCA 597
Qy 40 AspSerProValIleGluAlaValTyrThrAlaLeuLys 52
Db 598 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 636

RESULT 8
US-09-583-447A-1
; Sequence 1, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: WOJNOWSKI, Leszek
; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (105)..(1616)
US-09-583-447A-1
Alignment Scores:
Pred. No.: 0.853 Length: 1659
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
DB: 3 Gaps: 1

US-10-751-235-1 (1-77) x US-09-583-447A-1 (1-1659)
Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 582 CTGAGCGAGGAGCAGACAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 641
```

```

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThr--- 39
Db 642 ACCATGATGTAATCACTGCGCACATTTATTTGGAGTGAACCTTGGATTCTCTCAACAATCCA 701
Qy 40 AspSerProValIleGluAlaValTyrThrAlaLeuLys 52
Db 702 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 740

RESULT 9
US-09-583-447A-11
; Sequence 11, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: WOJNOWSKI, Leszek
; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1973
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(222)
; OTHER INFORMATION: n=a, c, t or g
US-09-583-447A-11
Alignment Scores:
Pred. No.: 1.09 Length: 1973
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
DB: 3 Gaps: 1

US-10-751-235-1 (1-77) x US-09-583-447A-11 (1-1973)
Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 940 CTGAGCGAGGAGCAGACAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 999
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThr--- 39
Db 1000 ACCATGATGTAATCACTGCGCACATTTATTTGGAGTGAACCTTGGATTCTCTCAACAATCCA 1059
Qy 40 AspSerProValIleGluAlaValTyrThrAlaLeuLys 52
Db 1060 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 1098

RESULT 10
US-08-802-141-3
; Sequence 3, Application US/08802141
; Patent No. 5773009
; GENERAL INFORMATION:
; APPLICANT: GLASS, ROGER I.
; APPLICANT: GENTSCH, JOHN R.
; APPLICANT: BHAN, M. K.
; APPLICANT: DAS, BIMAL K.
; TITLE OF INVENTION: ROTAVIRUS STRAIN AND RELATED
; TITLE OF INVENTION: COMPOSITIONS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
```

```

; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,141
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/231,041
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2352 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..2329
;
US-08-802-141-3
;
Alignment Scores:
Pred. No.: 1.66 Length: 2352
Score: 71.00 Matches: 21
Percent Similarity: 46.7% Conservative: 8
Best Local Similarity: 33.8% Mismatches: 23
Query Match: 18.0% Indels: 10
DB: 2 Gaps: 3

US-10-751-235-1 (1-77) x US-08-802-141-3 (1-2352)
;
QY 4 TyralaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMetThrLeuAsp 23
; Patent No. 6010703
;
Db 920 TATATGAGGCGATGCGATAGCTT-----GTGCTCATACTACTGTTCA 964
;
QY 24 ValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrThrAspSerProVal 43
;
Db 965 GTGAACGGTGTGTAGTGTATGATTATATACACTGGGTCGTTACCAACTGATTTCACGATC 1024
;
QY 44 IleGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuPro 63
;
Db 1025 CGAAAT-----TACACGTTTGTGAAGAAAGTTGCTTCGTATACGTTGAT----- 1069
;
QY 64 Tyrtip 65
;
Db 1070 TATTCG 1075
;
RESULT 11
US-08-911-321-6
; Sequence 6, Application US/08911321
; Patent No. 6010703
; GENERAL INFORMATION:
; APPLICANT: Roger K. Maes and Stephen J. Spatz
; TITLE OF INVENTION: Recombinant Forvirus
; TITLE OF INVENTION: Vaccine Against
; TITLE OF INVENTION: Feline Rhinotracheitis
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
;
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,183
; FILING DATE: July 26, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-166
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 6010703e
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6176
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Feline herpesvirus-1
; STRAIN: 1
; INDIVIDUAL ISOLATE: C-27
; CELL TYPE: N/A
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: Sequencing
; OTHER INFORMATION: DNA encoding PK, gG, gD,
; OTHER INFORMATION: gI, gE
;
US-08-911-321-6
;
Alignment Scores:
Pred. No.: 6.41 Length: 6176
Score: 71.00 Matches: 19
Percent Similarity: 45.61% Conservative: 7
Best Local Similarity: 33.3% Mismatches: 15
Query Match: 18.0% Indels: 16
DB: 3 Gaps: 2

US-10-751-235-1 (1-77) x US-08-911-321-6 (1-6176)
;
QY 3 ProTyraLalaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMetThr--- 21
;
Db 3668 CGGAAATGGAAGATGGTGGATCTACGCACCTGGCGGTAAAGATTACCATTAACAAA 3727
;
QY 22 LeuAspValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSer----- 36
;
Db 3728 GCTGATGATTGGCGCTTTCGTTGTTTACTCATTCGATACGCGTGCATCGACAT 3787
;
QY 37 -----LeuThrThrAspSerProVal 43
;
Db 3788 CATCGGACGAAATTTGAATGGTGAATTTCTTACTCTCCATCCCGATG 3838
;
RESULT 12
US-09-248-796A-4077
; Sequence 4077, Application US/09248796A
; Patent No. 6747137
```

GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4077
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4077

Alignment Scores:
Pred. No.: 0.261 Length: 432
Score: 69.50 Matches: 18
Percent Similarity: 53.97% Conservative: 16
Best Local Similarity: 28.57% Mismatches: 24
Query Match: 17.68% Indels: 5
DB: 3 Gaps: 2

US-10-751-235-1 (1-77) x US-09-248-796A-4077 (1-432)
QY 14 GluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPheAsn----- 31
DB 52 GATGCCAAATTAGTCAGGTTAATACAAATATCATTT-----TTATTTTCATGAAGAA 102
QY 32 TyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAlaLeu 51
DB 103 TATACCATTTACTGATATACCTAAAGACACACAAATTTTACAGTACTCGATTTCTAGCAAC 162
QY 52 LysGluAlaGluLeuArgSerThrAspLeuProTyrTyrIleAspAlaLeuCys 71
DB 163 AAAAAGGGCTCAATAAAGTTTCAGAGAGATATTTGCCCAAAATCAAAATAGACTTGGCGTGT 222
QY 72 LysIleVal 74
DB 223 ACCATTYA 231

RESULT 13
US-09-712-363-92
; Sequence 92, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-92

Alignment Scores:
Pred. No.: 3.08 Length: 2217
Score: 69.00 Matches: 21
Percent Similarity: 47.83% Conservative: 12
Best Local Similarity: 30.43% Mismatches: 25
Query Match: 17.56% Indels: 11
DB: 3 Gaps: 2

US-10-751-235-1 (1-77) x US-09-712-363-92 (1-2217)
QY 6 GluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIle 25
DB 301 GACGATGGCGGACGGCTTG-----CGGCTCTTCTGAGTGCAGCTGTCG----- 345
QY 26 GlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGlu 45
DB 346 -----GGCTACTCCATCGAGGACGTACTACTGCAGGACTCTCTGCTCGAT 390
QY 46 AlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuPro-TyrTr 65
DB 391 GCGGTCGAGGAGCGCTGCTGACCTGTCACCGAATCGCGGACCTGTACTCTGCTGAC 450
QY 65 LysIleAspAlaLeuCysValIle 73
DB 451 GGGTCCGGGGTCCCGCTGGCGCATC 475

RESULT 14
US-07-603-133B-19
; Sequence 19, Application US/07603133B
; Patent No. 5298244
; GENERAL INFORMATION:
; APPLICANT: Redmond, Mark J.
; APPLICANT: Ijaz, Mohammed K.
; APPLICANT: Parker, Michael D.
; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR
; USE IN A VACCINE TO ROTAVIRAL DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/603,133B
; FILING DATE: 19901025
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9313-0004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 2363 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..2337
US-07-603-133B-19

Alignment Scores:
Pred. No.: 3.37 Length: 2363
Score: 69.00 Matches: 21
Percent Similarity: 45.16% Conservatives: 7
Best Local Similarity: 33.87% Mismatches: 24
Query Match: 17.56% Indels: 10
DB: 2 Gaps: 3

US-10-751-235-1 (1-77) x US-07-603-133B-19 (1-2363)

QY 4 TyrAlaGluAspGlySerAlaValAlaAsnMetGluAlaLysPheSerGlnMetThrLeuAsp 23
Db 922 TATACAGAGATGGTGAAGAAGTTACT-----GCACATACTACGTGTTC A 966
QY 24 ValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProVal 43
Db 967 GTAATGGGAATAAATGATTTTAATTATATATATGTTGGATCATACCGACTGAT-----TTC 1020
QY 44 IleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuPro 63
Db 1021 GTAATATCAAAATATGAAGTGAATTAGGAA-----AATCTTTGTGTATATAGAC 1071
QY 64 TyrTIP 65
Db 1072 TACTG3 1077

RESULT 15

US-08-653-740-2
; Sequence 2, Application US/08653740
; Patent No. 5792850
; GENERAL INFORMATION:
; APPLICANT: James W. Baumgartner
; APPLICANT: Donald C. Foster
; APPLICANT: Frank J. Grant
; APPLICANT: Cindy A. Sprecher
; TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/653,740
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31,648
; REFERENCE/DOCKET NUMBER: 95-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2368 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23..1759
US-08-653-740-2

Alignment Scores:

Pred. No.: 3.38 Length: 2368
Score: 69.00 Matches: 24
Percent Similarity: 47.37% Conservatives: 12
Best Local Similarity: 31.58% Mismatches: 33
Query Match: 17.56% Indels: 7
DB: 2 Gaps: 2

US-10-751-235-1 (1-77) x US-08-653-740-2 (1-2368)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaAsnMetGluAlaLysPheSerGlnMet 20
Db 1940 CTTGGCCCCCCTGGGGGAAGACACACGGATGGAGTGCAGCAAGGAAATACATGAAATTC 1999
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsn-----Phe-AspSerLe 37
Db 2000 AGAGTGGCAGCTGCTGCCAAATCTGTCGCTGTAAACAGAACTGAATTTGGACCCCG 2059
QY 37 uThrThrAspSerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuAr 57
Db 2060 CCAGTGGCTCAGCCTGTATCCAGCACTTTGGCAGGCCAGGTGGAGGATCACTTAG 2119
QY 57 gSerThrAspLeuProTyrTyrLysIleAspAlaLeuCysLys 72
Db 2120 AGCTAGGAGTTTGAGACCAGCCTGG-----GCAATATGCAAG 2156

Search completed: January 15, 2006, 15:14:55
Job time : 177 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 06:11:50 ; Search time 3032 Seconds

(without alignments)
1188.194 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYAEDSANNWEAKPSQM.....STDLPLPYMKIDALCKIVPRQ 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10751235/runat_12012006_120915_2527/app_query.fasta_1.263
-DB=BST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LDOFCL=0 -LDOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235@cgn_1_1_5315@runat_12012006_120915_2527 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	1664	CNS0A5YJ	BX822288 Arabidops
2	393	100.0	1709	CNS0A665	BX822785 Arabidops
3	356	90.6	858	CO072553	CO072553 GR_Ea31M
4	346	88.0	736	CW839414	CW839414 GT8284.D8
5	343	87.3	531	BE552887	BE552887 946086H10
6	341	86.8	914	CK290930	CK290930 EST753644
7	337	85.8	840	BG643819	BG643819 EST512013

8	335	85.2	480	1	AJ486563	AJ486563
9	335	85.2	534	6	CA742365	CA742365 wflc.pk0
10	335	85.2	596	1	AJ432371	AJ432371
11	335	85.2	600	5	BU987393	BU987393 HF1K04F
12	329.5	83.8	570	6	BH584135	BH584135 BOHAD90TF
13	327	83.2	657	6	CB255014	54-E01836
14	326	83.0	668	8	DR952300	EST114383
15	326	83.0	886	8	DR923995	EST11553
16	324	82.4	469	8	CX541939	BL3DNF08H
17	324	82.4	652	2	BI263819	NF114G06P
18	320.5	81.6	708	9	BH931006	odil10g10.
19	319	81.2	626	5	BQ971938	OHB9C30.Y
20	317	80.7	897	8	CV886111	UCRCS04.2
21	307	78.1	693	10	CW166683	104_575_1
22	295	75.1	704	10	CZ710414	OC_Ba003
23	292	74.3	650	10	CZ799589	OC_Ba017
24	291	74.0	689	5	BQ862275	QGC20K13.
25	238	60.6	357	1	AJ615653	AJ615653
26	231	58.8	902	9	BZ822133	PUGG89TD
27	229	58.3	720	9	CC618386	OGVAP09TH
28	229	58.3	891	10	CG056732	CG056732 PUFTX10TB
29	227	57.8	662	9	CC394464	PUHF15TD
30	225	57.3	932	10	CG056734	CG056734
31	224	57.0	229	3	BI974897	gai74d11
32	224	57.0	1042	10	CW000851	ZMHBH9000
33	218	55.5	860	10	CZ321241	ZMHBH0023
34	214	54.5	679	7	CV538441	POD_005_F
35	210	53.4	598	3	BM003139	BM003139 1031108G0
36	207	52.7	474	3	BJ480615	BJ480615
37	207	52.7	598	3	RJ234910	RJ234910
38	207	52.7	634	8	CX023870	Wdlv4-404
39	207	52.7	636	1	AV941342	AV941342 AV941342
40	207	52.7	697	1	AV939356	AV939356
41	205	52.2	495	10	CG429086	CG429086 ZMHBH024
42	205	52.2	571	3	BJ471574	BJ471574 BJ471574
43	205	52.2	700	6	CD053804	CD053804
44	203	51.7	442	10	CG886172	CG886172 ZMHBH050
45	201	51.1	905	10	CL426941	ZMHBH044

ALIGNMENTS

RESULT 1

CNS0A5YJ

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSJTFB26Z804 of Flowers and buds of strain col-0 of Arabidopsis

thaliana (thale cress).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction ; Temple G.

CNS0A5YJ 1664 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSJTFB26Z804 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).

BX822288 GI:42464384

BX822288.1 GI:42464384

HTC; GSLT cDNA.

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 1664)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction ; Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source

1..1664

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="GSLTFB262B04"

/tissue_type="Flowers and buds"

/ecotype="Col-0"

/plasmid="pCMVSPORT_6"

1..1664

/gene="At3g53130"

ORIGIN

Alignment Scores:
Pred. No.: 3,248-45 Length: 1664
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-751-235-1 (1-77) x CNS0A5YJ (1-1664)

QY 1 LeuGinProTyraAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20

DB 624 TTGCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGGAAGCGAGTTCTCTCAGATG 683

QY 21 ThrLeuAppValileGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrAsp 40

DB 684 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAATTCGATTCCTTGACTACTGAT 743

QY 41 SerProValileGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60

DB 744 AGTCCTGTCTATGGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 803

QY 61 LeuLeuProTyTriPylsileAspAlaLeuCysLysIleValProArgGln 77

DB 804 CTTCTGCCATATGGAGATCGATGCAATTTGTGTAGATAGTCCCGAGACAG 854

RESULT 2

CNS0A665

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSLTFB64ZA07 of Flowers and buds of strain col-0 of Arabidopsis

thaliana (Chale cress).

EX822785

EX822785.1 GI:42464524

HTC; GSLT cDNA.

KEYWORDS

Arabidopsis thaliana

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1709)

Castelli,V., Aury J.M., Jaillon,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,P., Scarpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 1709)

Genoscope.

TITLE

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

COMMENT

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

Genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

[http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)

[length](http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

Location/Qualifiers

1..1709

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="GSLTFB64ZA07"

/tissue_type="Flowers and buds"

/ecotype="Col-0"

/plasmid="pCMVSPORT_6"

1..1709

/gene="At3g53130"

ORIGIN

Alignment Scores:
Pred. No.: 3,358-45 Length: 1709
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-751-235-1 (1-77) x CNS0A665 (1-1709)

QY 1 LeuGinProTyraAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20

DB 613 TTGCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGGAAGCGAGTTCTCTCAGATG 672

QY 21 ThrLeuAppValileGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrAsp 40

DB 673 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAATTCGATTCCTTGACTACTGAT 732

QY 41 SerProValileGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60

DB 733 AGTCCTGTCTATGGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 792

QY 61 LeuLeuProTyTriPylsileAspAlaLeuCysLysIleValProArgGln 77

DB 793 CTTCTGCCATATGGAGATCGATGCAATTTGTGTAGATAGTCCCGAGACAG 843

RESULT 3

CO072553

LOCUS

DEFINITION

GR_Ea31M14.r GR_Ea Gossypium raimondii cDNA clone GR_Ea31M14 3',

mRNA sequence.

CO072553

CO072553.1 GI:48742034

EST.

Gossypium raimondii

Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 858)

Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,

Udall,J.A., Rapp,R.A., Mendel,J.F., Rao,K., Soderlund,C. and

Wing,R.A.

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
SOURCE

```

FEATURES
  source
    Plate: 946086  Row: H  Column: 10.
    Location/Qualifiers
      1. .531
         /organism="Zea mays"
         /mol_type="mRNA"
         /cultivar="OH43"
         /db_xref="taxon:4577"
         /tissue_type="tassels"
         /dev_stage="just after the transition from vegetative to
         inflorescence development"
         /lab_host="XL0Lg"
         /clone_lib="946 - tassel primordium prepared by Schmidt
         lab"
         /note="Organ: tassels; Vector: HybrizAP; Site_1: ECOR1;
         Site_2: XhoI; George Chuck dissected immature tassels
         between 1mm and 3mm. Sharon Stanfield prepared the cDNA
         library in HybrizAP. Sample insert size range was 350 bp
         to 3 kb with a 1 Kb average."

```

ORIGIN		
Alignment Scores:		
Pred. No.:	1.12e-38	531
Score:	343.00	65
Percent Similarity:	92.21%	Conservative:
Best Local Similarity:	84.42%	Mismatches:
Query Match:	87.28%	Indels:
DR:	2	Gaps:
		0

US-10-751-235-1 (1-77) x BE552887 (1-531)

Qy	1	LeuGlnProTyrTrpAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet	20
Db	1	CTTGAGCCATATGCTTTGAGTGGGGAACCTGCTCAATATGGAAGCAGAGTTTCTCAGTTG	60
Qy	21	ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp	40
Db	61	ACATTGATGTGATGTGTTTATCATCTGTTCAACTACAAATTTTGATTTCCCTCACAAACAGAT	120
Qy	41	SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp	60
Db	121	AGTCCTGTCAATGTATGCTGTTTACTTGCACTCAAGAGCAGAGCTTCGTCTCTACAGAT	180
Qy	61	LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln	77
Db	181	CTTTTGGCATACTGCAAGTGGTGTCTTTGTGCAAGATAATCCCAAGACAG	231

RESULT 6	CK290930	914 bp	mRNA	linear	EST 02-AUG-2004
LOCUS	CK290930				
DEFINITION	EST573644 Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NEMBY36 5' end. mRNA sequence.				

ACCESSION	CK290930
VERSION	CK290930.1
KEYWORDS	GI:39870877
SOURCE	EST.
ORGANISM	<i>Nicotiana benthamiana</i>
	<i>Nicotiana benthamiana</i>

ORGANISM

ascellus, tamulus, cotinatus, notatus;
1 (bases 1 to 914)

REFERENCE
AUTHORS
1. (PAGES 1 TO 21)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

Staskawicz, B., Jin, H. and Baker, B.

TITLE
Generation of EST sequences from *Nicotiana benthamiana*

JOURNAL Unpublished (2003)

COMMENT	Other_ESTS: EST753645

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

FEATURES
Location/Qualifiers
1. .914

```

1. .914
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMEY36"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMWsport6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN	Alignment Scores:	Length:	914
Pred. No.:	4,438-38	Matches:	66
Score:	341.00	Conservative:	6
Percent Similarity:	93.51%	Mismatches:	5
Best Local Similarity:	85.71%	Indels:	0
Query Match:	86.77%	Gaps:	0
DB:	7		

UUS-10-751-235-1 (1-77) x CK290930 (1-914)

Qy	1	LeuGlnPro	Trp	Trp	Ala	Glu	Asp	Gly	Ser	Ala	Val	Asn	Met	Glu	Ala	Leu	Phs	Ser	Gln	Met	20	
Db	481	CTT	AC	AC	CT	GC	AT	TT	CT	GG	CT	CT	GC	AG	TA	AT	TGG	AGG	CA	AGG	TTT	CT
Qy	21	Thr	Leu	Asp	Val	Ile	Gly	Leu	Ser	Leu	Phs	Asn	Trp	Asn	Phs	Asp	Ser	Leu	Thr	Thr	Asp	
Db	541	AC	ACT	TG	AT	TG	GT	CT	CT	CT	CA	AT	TT	GAT	TT	CC	CT	TACT	ACT	TG	AC	
Qy	41	Ser	Pro	Val	Ile	Glu	Ala	Val	Thr	Ala	Leu	Lys	Glu	Ala	Glu	Leu	Arg	Ser	Thr	Asp	60	
Db	601	AGT	CC	AGT	ATT	TG	AG	CA	GTT	AT	CAC	TG	CGT	TAA	AG	AG	CAG	CA	ACT	CCG	TT	
Qy	61	Leu	Leu	Pro	Trp	Trp	Lys	Ile	Asp	Ala	Leu	Cys	Lys	Ile	Val	Pro	Arg	Gln			77	
Db	561	CTG	TTT	CC	CT	ATT	TGC	AG	CT	CA	AG	CC	TAT	TGT	AGG	TC	AT	CC	CA	CA	711	

DEPT. 7

RESULT /
RG643819

51001
51002
51003

LOCUS DEFINITION

DEFINITIONS

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANIS

REFERENCE

AUTHORS

DATE

TITLE

**JOURNAL
COMMENT**

COMMENT:


```

Score:          335.00      Matches:          65
Percent Similarity: 90.91%      Conservative: 5
Best Local Similarity: 84.42%      Mismatches: 7
Query Match:      85.24%      Indels: 0
DB:               6          Gaps: 0

US-10-751-235-1 (1-77) x CA742365 (1-534)

Qy 1 LeuGlnProTyrrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 82 CTCGAGACTTATGCTTTCGAGCGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAAATG 141
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrThrAsp 40
Db 142 ACATTAGATGTGATTGGTTATCCCTTGTTCACACTCAACTCAACTTGTGATTCCTCCATCAGAT 201
Qy 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 202 AGTCCTGTTATGATGCTTGTACACTGCATCAAGAAGCTGAGGCTCGTTCACAGAT 261
Qy 61 LeuLeuProTyrrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 262 CTTTACCATACTGGCAGATCGATTGCTGTGCAAGATTGTTCTCTAGACAG 312

RESULT 10
AJ432371 596 bp mRNA linear EST 15-MAR-2002
LOCUS AJ432371 800011 Hordeum vulgare cDNA clone S0001100019A01F1, mRNA
DEFINITION AJ432371 sequence.
ACCESSION AJ432371
VERSION AJ432371.1 GI:19520823
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 596)
Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
Barley EST's
Unpublished (2002)
Contact: Schulman AH
Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
Finland.

FEATURES
source
1..596
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
/clone="S0001100019A01F1"
/dev_stage="Developing seed"
/clone_lib="S00011"
/note="12,15,18 days after pollination"

ORIGIN
Alignment Scores:
Pred. No.: 1 84e-37 Length: 596
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 1 Gaps: 0

US-10-751-235-1 (1-77) x AJ432371 (1-596)

Qy 1 LeuGlnProTyrrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 226 CTCGAGACATATGCTTTCGAGCGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAAATG 285
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrThrAsp 40

```

```

Db 286 ACACATAGATGTGATTGGTTTGTCTTGTTCACATACACTTGTGATTCCTCCATCAGAT 345
Qy 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 346 AGTCCTGTTATGATGCTTGTACACCGCACTGAAGAAGCAGAGGCTCGTTCACAGAT 405
Qy 61 LeuLeuProTyrrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 406 CTTTACCATACTGGCAGATTGATTGCTGTGCAAGATTGTTCTCTAGACAG 456

RESULT 11
BU987393 600 bp mRNA linear EST 22-OCT-2002
LOCUS BU987393 5-PRIME, mRNA sequence.
DEFINITION BU987393
ACCESSION BU987393.1 GI:24238339
VERSION BU987393
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 600)
Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.
Barley ESTs from developing seeds
Unpublished (2002)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 600 Std Error: 0.00
Plate: 14 row: K column: 4
Seq primer: M13rev.

FEATURES
source
1..600
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Barke"
/sub_species="vulgare"
/db_xref="GABI:247336"
/db_xref="taxon:112509"
/clone="HF14K04"
/tissue_type="caryopsis"
/dev_stage="developing caryopsis, 16-25 DAF (days after
flowering)"
/lab_host="XL10-Gold"
/clone_lib="HP"
/note="Vector: pBluescript SK+; Site1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); developing
caryopsis, 16-25 DAF(days after flowing) Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRIadapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI,PstI). NOTE: Also
due to the cloning system used Blue/White selection for
recombinants is not 100% reliable.Average insert size is
940 bp"

ORIGIN
Alignment Scores:
Pred. No.: 1 85e-37 Length: 600
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 5 Gaps: 0

US-10-751-235-1 (1-77) x BU987393 (1-600)

```

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 |||:::|||||
 Db 176 CTCGACACATATGCTTTGAGCGGTGAACCTGTTAATATGGAAGCGAGATTTTCTCAATG 235
 |||:::|||||
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 |||:::|||||
 Db 236 ACACATAGATGATGTTGTTGCTTTGTTCACTACACTTGTGATTCCTCAGATCAGAT 295
 |||:::|||||
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 |||:::|||||
 Db 296 AGTCCTGTTATTGATGCTGTTTACACCGCACTGMAAGAGCAGAGGCTGTTCTACAGAT 355
 |||:::|||||
 QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
 |||:::|||||
 Db 356 CTTTACCATACTGGCAGATGATTGCTGTGCAAGATTGTTCTCAGACAG 406
 |||:::|||||

RESULT 12
 BH584135/c
 LOCUS BH584135
 DEFINITION BOHAD90TF BOHA Brassica oleracea genomic clone BOHAD90, genomic survey sequence.
 ACCESSION BH584135
 VERSION BH584135.1 GI:17836592
 KEYWORDS GSS
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 570)
 AUTHORS Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
 TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
 JOURNAL Genome Res. 15 (4), 487-495 (2005)
 PUBMED 15805490
 COMMENT Other_GSSs: BOHAD90TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TP
 Class: sheared ends.

FEATURES
 source
 1..570
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T0100DH3"
 /db_xref="taxon:3712"
 /clone="BOHAD90"
 /clone_lib="BOHA"
 /note="Vector: pHOS1; Site 1: BatXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.07e-36 Length: 570
 Score: 329.50 Matches: 71
 Percent Similarity: 71.84% Conservatve: 3
 Best Local Similarity: 68.93% Mismatches: 3
 Query Match: 83.84% Indels: 26
 DB: 9 Gaps: 1

US-10-751-235-1 (1-77) x BH584135 (1-570)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 |||:::|||||
 Db 442 CTCGACCTTATGATGAGTGGAGAAAGCTGTGACATGAGAGAGATTTCTCAGTTA 383
 |||:::|||||
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 |||:::|||||

Db 382 ACTCTCGAGCGATGGCTTATCTCTTTTAACTACAACCTTTGATTCCTACTACCGAT 323
 |||:::|||||
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 |||:::|||||
 Db 322 AGTCCTGTCATCGAAGCTGTTTACACTGCTCTCAAGAAGCTGAGCTTGCTTACTAGT 263
 |||:::|||||
 QY 61 LeuLeuProTyrTrpLys-----
 |||:::|||||
 Db 262 ATTCTACCCCTATTGGAAGCGCAAGTTTGTGTTCTTCTGTGGCGCAGACTTGGATTCT 203
 |||:::|||||
 QY 67 -----
 |||:::|||||
 Db 202 TCTTGAGCCTTGGTGTGTTTCTTGTTCAGATCGATCGGTGTGTGAAGATAGTTCC 143
 |||:::|||||
 QY 75 oArgGln 77
 |||:::|||||
 Db 142 GAGACAA 136
 |||:::|||||
 RESULT 13
 CB255014
 LOCUS CB255014
 DEFINITION 54-E018363-019-007-L14-T7R MP1Z-ADIS-019 Arabidopsis thaliana cDNA clone MP1Zp768L147Q 5-PRIME, mRNA sequence.
 ACCESSION CB255014
 VERSION CB255014.1 GI:56915639
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 657)
 AUTHORS Jakoby,M., Stracke,R., Soerensen,T.R. and Weisshaar,B.
 TITLE Arabidopsis thaliana cDNA library enriched in transcription factors
 JOURNAL Unpublished (2003)
 COMMENT Contact: Weisshaar B
 ADIS DNA core facility at MP1Z
 Max-Planck-Institute for Plant Breeding Research
 Carl-von-Linne Weg 10, 50829 Koeln, Germany
 Fax: 00492215062851
 Email: weisshaar@mpiz-koeln.mpg.de
 Insert Length: 657 Std Error: 0.00
 Plate: 7 row: L column: 14
 Seq primer: T7R; CTAATACGACTCACTATAGGGA.
 FEATURES
 Location/Qualifiers
 1..657
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultiivar="At7"
 /db_xref="GABI:597448"
 /db_xref="taxon:3702"
 /clone="MP1Zp768L147Q"
 /tissue_type="hypocotyl"
 /dev_stages="tissue culture"
 /lab_host="E. coli DH5alpha mcr"
 /clone_lib="MP1Z-ADIS-019"
 /note="Vector: pSPORT1; Site 1: NotI primer adapter; Site 2: SalI primer adapter; RNA from cellculture (At7) 5 days after inoculation treated with 0.002 mM cycloheximide for 2 h in the dark. Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN
 Alignment Scores:
 Pred. No.: 2.96e-36 Length: 657
 Score: 327.00 Matches: 68
 Percent Similarity: 97.18% Conservatve: 1
 Best Local Similarity: 95.77% Mismatches: 0
 Query Match: 83.21% Indels: 2
 DB: 6 Gaps: 0

US-10-751-235-1 (1-77) x CB255014 (1-657)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
 Db 447 TTGCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGGAAGCAAGTCTCTCAGATG 506
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAenTyrAenPheAenPheSerLeuThrThrAsp 40
 Db 507 ACACCTGATGTCATGGGTGTCTCTTTTAACTACAAATTCGATTCCTTGACACTACTGAT 566
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAla--GluLeuArgSerThrA 60
 Db 567 AGTCCTGTCATGAAGCTGTGTACACTGCTCTTAAAGAAGCTTGAGCTTGCTTCTACTG 626
 QY 60 spLeuLeuProTyrTrpLysIleAspAla 69
 Db 627 ATCTTCTGCCATATTGGAAGATCGATGCA 655
 DR952300 668 bp mRNA linear EST 02-AUG-2005
 EST1143839 Aquilegia cDNA library Aquilegia formosa x Aquilegia
 pubescens cDNA clone COLS013, mRNA sequence.
 DR952300
 DR952300.1 GI:71721663
 EST.
 Aquilegia formosa x Aquilegia pubescens
 Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 1 (bases 1 to 668)
 Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
 Nordborg, M. and Tomkins, J.
 Generation of ESTs from Aquilegia
 Unpublished (2005)
 Other ESTs: EST1143838
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: M13 Reverse.

FEATURES
 source

1..668
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="COLS013"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by
 weight respectively) and 3) Shoot apical meristems. A
 fourth set of tissue was collected from plants of A.
 formosa. These plants were grown from seed in sand and at
 approximately 1 month root tissue and leaf tissue of
 various developmental stages were collected (84 & 16% by
 weight respectively). Total RNA was extracted from each
 set of tissue and pooled in the following proportions:
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
 total RNA, mRNA was extracted and enriched for full-length
 messages and then normalized with proprietary methods by

Invitrogen."

ORIGIN

Alignment Scores: Length: 668
 Pred. No.: 4.21e-36 Matches: 63
 Score: 326.00 Conservative: 6
 Percent Similarity: 89.61% Mismatches: 8
 Best Local Similarity: 81.82% Indels: 0
 Query Match: 82.95% Gaps: 0
 Db: 8

US-10-751-235-1 (1-77) x DR952300 (1-668)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
 Db 428 CTTAGACCTGATGCTGTGAGTGGTCTCTGTGAATATGGAAGCAAGTCTTCGAGTTA 487
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAenTyrAenPheAenPheSerLeuThrThrAsp 40
 Db 488 ACTTAGATGTTATAGGTTTGTCACTGTTCAATATATATTTTGAATTCATTAAATCGGAT 547
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 Db 548 AGTCCTGTTATGATGTCAGTATATCTGCTTTAAAGAGGCGAGGCTCGCTCAACAGAT 607
 QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
 Db 608 CTTTACCATATTGGAAGATTAACTTTTGTGCAAGATAATTCGAGGCA 658

RESULT 15

DR923995

LOCUS

DEFINITION

Pubescens cDNA clone COLM240, mRNA sequence.

ACCESSION

DR923995

VERSION

DR923995.1

KEYWORDS

SOURCE

ORGANISM

Aquilegia formosa x Aquilegia pubescens

Aquilegia formosa x Aquilegia pubescens

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;

Ranunculaceae; Aquilegia.

1 (bases 1 to 886)

Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,

Nordborg, M. and Tomkins, J.

Generation of ESTs from Aquilegia

Unpublished (2005)

Other ESTs: EST115533

Contact: Scott Hodges

Department of Ecology, Evolution and Marine Biology

University of California, Santa Barbara

Santa Barbara, CA 93106, USA

Tel: 805 893 7813

Fax: 805 893 4724

Email: hodges@lifesci.ucsb.edu

Seq primer: M13 Reverse.

Location/Qualifiers

1..886

/organism="Aquilegia formosa x Aquilegia pubescens"

/mol_type="mRNA"

/db_xref="taxon:338618"

/clone="COLM240"

/tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"

/lab_host="DH10B T1 (T1 and T5 phage resistance)"

/clone_lib="Aquilegia cDNA library"

/note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by

weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.:	6.12e-36	Length:	886
Score:	326.00	Matches:	63
Percent Similarity:	89.61%	Conservative:	6
Best Local Similarity:	81.82%	Mismatches:	8
Query Match:	82.95%	Indels:	0
DB:	8	Gaps:	0

US-10-751-235-1 (1-77) x DR923995 (1-886)

Qy	1	LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet	20
Db	646	CTTAGACCTGATGCTGTGAGTGGTCTCTGTGAATATGGAGGCAAGTTTTCGCAGTTA	705
Qy	21	ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp	40
Db	706	ACTTTAGATGTTATAGGTTTGTCACCTTCANTATATTTTGTATTCACCTTAATCGCGAT	765
Qy	41	SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp	60
Db	766	AGTCCTGTTATTGATGCAGTATATCTGCATTAAAAGAGGACAGGCTCGCTCAACAGAT	825
Qy	61	LeuLeuProTyrTrpLysIleaspAlaLeuCysLysIleValProArgGln	77
Db	826	CTTCTACCATATTGGAAGATTAACTTTTGTGCAAGATAATTCCGAGGCAA	876

Search completed: January 15, 2006, 15:11:57
Job time : 3038 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 13:12:22 ; Search time 615 Seconds
(without alignment)
1035.353 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQYABDGSANWEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10751235/runat 12012006 120916 2564/app query.fasta_1.263
-DB=Published Applications NA Main -QFWT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235.ecgn 1 1026 @runat 12012006 120916 2564 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESOURCY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -THREADS=6 -DELEXT=7

Database :

Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	3377	8	US-10-739-930-773
2	346	88.0	2467	9	US-10-751-235-5
3	346	88.0	2467	9	US-10-751-235-7
4	346	88.0	4170	9	US-10-751-235-6
5	343	87.3	531	9	US-10-751-235-26
6	343	87.3	1804	7	US-10-751-235-114-32608
7	343	87.3	2537	8	US-10-425-115-148190

ALIGNMENTS

RESULT 1

US-10-739-930-773
; Sequence 773, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 773
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER3001_1
US-10-739-930-773

Alignment Scores:

Pred. No.:	8.9e-51	Length:	3377
Score:	393.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-10-751-235-1 (1-77) x US-10-739-930-773 (1-3377)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20

Db 676 TTGCAGCCTTATGCAGACGAGAGTCTGTGATATATGAGAGCGAAGTTCTCTCAGATG 735
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 736 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 795
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 796 AGTCCTGTCATTTGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 855
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysValLysIleValProArgGln 77
Db 856 CTTCTGCCATATTTGAAGATCGATGCAATTTGTGTAAAGATAGTCCCGAGACAG 906

RESULT 2

US-10-751-235-5
; Sequence 5, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-5

Alignment Scores:
Pred. No.: 1.65e-43 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-5 (1-2467)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 747 TTGCAGCCTTATGCAGACGAGAGTCTGTGATATATGAGAGCGAAGTTCTCTCAGATG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 807 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 866
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTCATTTGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 926
Qy 61 LeuLeuProTyrTrpLysIle-----
Db 927 CTTCTGCCATATTTGAAGAGC-AAGTTTCTGTGTTTTTCTGTGGTTTGTGATTGTGTG 985
Qy 68 -----AspAlaLeu 70
Db 986 GAACAATTGGAATCTTGTAAATTGAGAGGTTTGGTTGTTTTTTCAGATCGATGCAATG 1045
Qy 71 CysLysIleValProArgGln 77
Db 1046 TGTAAAGATAGTCCCGAGACAG 1066

RESULT 3

US-10-751-235-7
; Sequence 7, Application US/10751235

; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-7

Alignment Scores:
Pred. No.: 1.65e-43 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-7 (1-2467)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 747 TTGCAGCCTTATGCAGACGAGAGTCTGTGATATATGAGAGCGAAGTTCTCTCAGATG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 807 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 866
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTCATTTGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 926
Qy 61 LeuLeuProTyrTrpLysIle-----
Db 927 CTTCTGCCATATTTGAAGAGC-AAGTTTCTGTGTTTTTCTGTGGTTTGTGATTGTGTG 985
Qy 68 -----AspAlaLeu 70
Db 986 GAACAATTGGAATCTTGTAAATTGAGAGGTTTGGTTGTTTTTTCAGATCGATGCAATG 1045
Qy 71 CysLysIleValProArgGln 77
Db 1046 TGTAAAGATAGTCCCGAGACAG 1066

RESULT 4

US-10-751-235-6
; Sequence 6, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 4170
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-6

Alignment Scores:
Pred. No.: 3.6e-43 Length: 4170
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: 9 Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-6 (1-4170)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 1756 TTGCAGCCTTATGCAGAGAGGAGTGTGTGAATATGGAAGCGAGTCTCTCAGATG 1815
Qy 21 ThrLeuaspValIleGlyLeuSerLeuPheAenTyrAenPheAenSerLeuThrAsp 40
Db 1816 ACATGTGATGTCATGGGTGTCTCTTTTAACTACAAATTTGATTTCTTGACTACTGAT 1875
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 1876 AGTCTGTGTCATGAAGCTGTGTACACTGCTCTTAAAGAAGCTGAGCTTCTGTTCTACTGAT 1935
Qy 61 LeuLeuProTyrTrpLysIle----- 67
Db 1936 CTTTGCCCATATGGAAGGC-AGGTTTCTGTGTTTTTTCTGTGTTTGTGATTTGTG 1994
Qy 68 -----AspAlaLeu 70
Db 1995 GAACAATTTGGATCTTGTTAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGATTG 2054
Qy 71 CysLeuValIleValProArgGln 77
Db 2055 TGTAAGATAGTCCCGAGACAG 2075

RESULT 5
US-10-751-235-26
; Sequence 26, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Zea mays
US-10-751-235-26

Alignment Scores:
Pred. No.: 5.05e-44 Length: 531
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 0
Query Match: 87.28% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-26 (1-531)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 1 CTTGAGCCATATGCTTTGAGTGGGGAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTG 60
Qy 21 ThrLeuaspValIleGlyLeuSerLeuPheAenTyrAenPheAenSerLeuThrAsp 40
Db 61 ACATGTGATGTTGTTTATCATTTGTTCAACTACAAATTTTGTATTCCTCTCAGACAGAT 120

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 121 AGTCTGTGTCATGATGCTGTTTATCTGCACTCAAGAAGCAGAGCTTCTGTTCTACAGAT 180
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 181 CTTTGGCCATATGGAAGGTTGGTTTCTTGTGCAAGATAATCCCAAGACAG 231

RESULT 6

US-10-425-114-32608
; Sequence 32608, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32608
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17009E12_FLI
US-10-425-114-32608

Alignment Scores:
Pred. No.: 3.1e-43 Length: 1804
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 0
Query Match: 87.28% Indels: 0
DB: 7 Gaps: 0

US-10-751-235-1 (1-77) x US-10-425-114-32608 (1-1804)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 494 CTTGAGCCATATGCTTTGAGTGGGGAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTG 553
Qy 21 ThrLeuaspValIleGlyLeuSerLeuPheAenTyrAenPheAenSerLeuThrAsp 40
Db 554 ACATGTGATGTTGTTTATCATTTGTTCAACTACAAATTTTGATTCCTCTCAGACAGAT 613
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 614 AGTCTGTGTCATGATGCTGTTTATCTGCACTCAAGAAGCAGAGCTTCTGTTCTACAGAT 673
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 674 CTTTGGCCATATGGAAGGTTGGTTTCTTGTGCAAGATAATCCCAAGACAG 724

RESULT 7

US-10-425-115-148190
; Sequence 148190, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 148190
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MWT4577_66678C.1
US-10-425-115-148190

Alignment Scores:
Pred. No.: 5, 15e-43 Length: 2537
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 6
Query Match: 87.28% Indels: 0
DB: 8 Gaps: 0

US-10-751-235-1 (1-77) x US-10-425-115-148190 (1-2537)

Qy 1 LeuGlnProTyrrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 747 CTTCGACCATATGCTTTCAGTGGGGAACCTGTCAATATGGAAGCAGGCTTTCTCAGTTG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrAsp 40
Db 807 ACATTGGATGTGATTGGTTTATCATTTGTCAACTACAATTTTGTATCCCTCACACAGAT 866
Qy 41 SerProValIleGluAlaValTyrrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTCATGATGCTGTGTTTATCTGCACTCAAGAAGCAGAGCTTCGTTCTACAGAT 926
Qy 61 LeuLeuProTyrrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 927 CTTTTCACCATATCGAAGCTTGGTTCTGTGCAAGATAATCCCAAGACAG 977

RESULT 8

US-10-751-235-25
; Sequence 25, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 839
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-10-751-235-25

Alignment Scores:
Pred. No.: 8, 95e-43 Length: 839
Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-25 (1-839)

Qy 1 LeuGlnProTyrrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 31 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGAGGAGGAAAAGTTTCTCAACTA 90
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrAsp 40
Db 91 ACATTGGATGTATTGGCCCTTGCACTTCAATTACAATTTGATTCCTTACTACTGAC 150

Qy 41 SerProValIleGluAlaValTyrrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 151 AGTCAGTTATTGATGAGTTTACACTGCACATAAAGAGCAGAACTCCGTTCACTGAT 210
Qy 61 LeuLeuProTyrrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 211 TTGTTGCCATATTGGCAGATCAAGAGCTTTATGTAAGTTTCATCCCAAGACAA 261

RESULT 9

US-10-751-235-24
; Sequence 24, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-751-235-24

Alignment Scores:
Pred. No.: 2, 73e-42 Length: 1086
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-24 (1-1086)

Qy 1 LeuGlnProTyrrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 589 CTCGAGACTTATGCTTTCAGTGGTGAACCTGTTATATATGGAAGCAGGTTTCTCAATG 648
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrAsp 40
Db 649 ACATTAGATGTGATTGGTTTATCCTTGTTCACACTACAACCTTGTATCCCTCACATCAGAT 708
Qy 41 SerProValIleGluAlaValTyrrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 709 AGTCCTGTTTATGATGCTGTTTACACTGCACCTCAAGAAGCTGAGGCTCGTTCTACAGAT 768
Qy 61 LeuLeuProTyrrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 769 CTTTATACCATACTGGCAGATCGATTTGCTGTGCAAGATTGTTCTCTACAGAG 819

RESULT 10

US-10-437-963-15200
; Sequence 15200, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Alignment Scores:
Pred. No.:          Length:      1118
Score:              2,85e-42      66
Percent Similarity: 335.00        Matches:
Best Local Similarity: 90.91%      Conservative: 4
Query Match:         85.71%        Mismatches: 7
DB:                  Indels:       0
                       Gaps:       0

US-10-751-235-1 (1-77) x US-10-437-963-15200 (1-1118)

QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 166 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 225
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 226 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTTACATTTTGTCCCTCAGATCAGAT 285
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 286 AGCCCTGTTATGTGCTGTTTACACTGCACCTCAGAGGAGCAAGCACTTCGTTCTAAGAT 345
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 346 CTTTACCATACTGGAAGATGATTTGCTGTGCAAGATTTGTTCTTAGACAA 396

RESULT 11
US-10-751-235-23
; Sequence 23, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Hordeum vulgare
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t
US-10-751-235-23

Alignment Scores:
Pred. No.:          Length:      1638
Score:              5.03e-42      65
Percent Similarity: 335.00        Matches:
Best Local Similarity: 90.91%      Conservative: 5
Query Match:         84.42%        Mismatches: 7
DB:                  Indels:       0
                       Gaps:       0

US-10-751-235-1 (1-77) x US-10-751-235-23 (1-1638)

QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 166 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 225
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 226 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTTACATTTTGTCCCTCAGATCAGAT 285
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 286 AGCCCTGTTATGTGCTGTTTACACTGCACCTCAGAGGAGCAAGCACTTCGTTCTAAGAT 345
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 346 CTTTACCATACTGGAAGATGATTTGCTGTGCAAGATTTGTTCTTAGACAA 396

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)

QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTTACATTTTGTCCCTCAGATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTTATGTGCTGTTTACACTGCACCTCAGAGGAGCAAGCACTTCGTTCTAAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTTGCTGTGCAAGATTTGTTCTTAGACAA 852

RESULT 13
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
```

```
Db 574 CTCGAGACATATGCTTTGAGCGGTGAACCTGTAATATGGAAGCGAGATTTTCTCAATG 633
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 634 ACACTAGATGTGATGGTTGGTTGCTTGTTCACCACTTACCACTTTGATTCCCTCAGATCAGAT 693
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 694 AGTCCTGTTATGTGCTGTTTACACCGCACCTGGAAGAGCAGAGGCTCGTTCTACAGAT 753
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 754 CTTTACCATACTGGAAGATGATTTGCTGTGCAAGATTTGTTCTTAGACAG 804

RESULT 12
US-10-751-235-22
; Sequence 22, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-751-235-22

Alignment Scores:
Pred. No.:          Length:      1686
Score:              5.25e-42      66
Percent Similarity: 335.00        Matches:
Best Local Similarity: 90.91%      Conservative: 4
Query Match:         85.71%        Mismatches: 7
DB:                  Indels:       0
                       Gaps:       0

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)

QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTTACATTTTGTCCCTCAGATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTTATGTGCTGTTTACACTGCACCTCAGAGGAGCAAGCACTTCGTTCTAAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTTGCTGTGCAAGATTTGTTCTTAGACAA 852

RESULT 13
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
```

```
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-10-751-235-27

Alignment Scores:
Pred. No.: 4.18e-40 Length: 624
Score: 319.00 Matches: 60
Percent Similarity: 89.61% Conservative: 9
Best Local Similarity: 77.92% Mismatches: 8
Query Match: 81.17% Indels: 0
Gaps: 0
DB:

US-10-751-235-1 (1-77) x US-10-751-235-27 (1-624)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 253 CTAAGTCATACGCGAGTGCACAGCTGCTTAAACATGCGAGCAACAGTTTCGCAGTTA 312
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 313 ACCCTTGATGTTATCGGTCTAGCGGTATTTAACTACAATTTTGACTCTACTTACGCCCGAT 372
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 373 AGTCCTGTAATGATCTGTTATACCGCACTTAAAGAGCTGAAGCCGTTCAACTGAT 432
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 433 CTTTGGCCATATTGGAAGATAAGTCGGTTATGTGAAGATTATACCAAGACAA 483

RESULT 14
US-10-424-599-9451
; Sequence 9451, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9451
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108543C.1
US-10-424-599-9451

Alignment Scores:
Pred. No.: 1.19e-25 Length: 229
Score: 224.00 Matches: 46
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 70.77% Mismatches: 13
Query Match: 57.00% Indels: 0
Gaps: 0
DB:

US-10-751-235-1 (1-77) x US-10-424-599-9451 (1-229)
QY 2 GlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThr 21
Db 29 CRAACTGATGCACTTAATGGAAGTCTGCTGCTCAACATCGAGGCAAAAGTTCTCGCAGCTGACT 88
QY 22 LeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSer 41
```

```
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-10-751-235-27

Alignment Scores:
Pred. No.: 4.18e-40 Length: 624
Score: 319.00 Matches: 60
Percent Similarity: 89.61% Conservative: 9
Best Local Similarity: 77.92% Mismatches: 8
Query Match: 81.17% Indels: 0
Gaps: 0
DB:

US-10-751-235-1 (1-77) x US-10-751-235-27 (1-624)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 253 CTAAGTCATACGCGAGTGCACAGCTGCTTAAACATGCGAGCAACAGTTTCGCAGTTA 312
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 313 ACCCTTGATGTTATCGGTCTAGCGGTATTTAACTACAATTTTGACTCTACTTACGCCCGAT 372
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 373 AGTCCTGTAATGATCTGTTATACCGCACTTAAAGAGCTGAAGCCGTTCAACTGAT 432
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 433 CTTTGGCCATATTGGAAGATAAGTCGGTTATGTGAAGATTATACCAAGACAA 483

RESULT 14
US-10-424-599-9451
; Sequence 9451, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9451
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108543C.1
US-10-424-599-9451

Alignment Scores:
Pred. No.: 1.19e-25 Length: 229
Score: 224.00 Matches: 46
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 70.77% Mismatches: 13
Query Match: 57.00% Indels: 0
Gaps: 0
DB:

US-10-751-235-1 (1-77) x US-10-751-235-48 (1-1101)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 100 CTAGACAAGTATGCGCGCTCAGGCACCGCTGGACATGGAAAACCTTCTTCAGCCGGCTG 159
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 160 GGTCCTGGACATCATCGCAAGCGCTGTTCAACTACGACTCGCTCGCTGGCGCACGAC 219
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 220 GACCCCGTCATCCAGCGCGTGTACACGTTGCTGCGGGAAGCGGAGCACCCTCCACAGCG 279
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 280 CCCATCGCCTACTGGACATTTCCCGCATCCAGTTTGTGTGTCGCGCGGCAG 330

Search completed: January 15, 2006, 15:25:16
Job time : 619 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: January 15, 2006, 05:58:40 ; Search time 3652 Seconds
(without alignments)
1198.507 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYAEDGSANMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2 1/USPTO spool/US10751235/runat 12012006 120915 2519/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CNC 1.1 4939 @runat 12012006 120915 2519 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.in:*
3: gb.env:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pr:*
9: gb.ro:*
10: gb.scs:*
11: gb.sy:*
12: gb.un:*
13: gb.vi:*
14: gb.htg:*
15: gb.pli:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	1620	15	AY424805 Arabidops
2	393	100.0	1858	15	AY091083 Arabidops
3	393	100.0	3304	15	AF367289 Arabidops

4	346	88.0	92611	15	ATT4D2	AL132958 Arabidops
5	337	85.8	1867	15	BT012891	BT012891 Lycopersi
6	335	85.2	1876	15	AK065689	AK065689 Oryza sat
7	303	77.1	110000	15	AP008216_208	Continuation (209
8	303	77.1	178022	15	AC025783	AC025783 Oryza sat
9	303	77.1	299735	15	AB017117	AB017117 Oryza sat
10	227	57.8	205408	14	AC155564	AC155564 Zea mays
11	220	56.0	110000	14	CR954213_4	Continuation (5 of
12	211	53.7	110000	14	CR954201_08	Continuation (9 of
13	206	52.4	4217	15	AK068163	AK068163 Oryza sat
14	206	52.4	110000	14	CR954209_3	Continuation (4 of
15	201	51.1	2038	15	AY601887	AY601887 Ginkgo bi
16	200	50.9	1831	15	AF022457	AF022457 Glycine m
17	199	50.6	2093	15	AK100596	AK100596 Oryza sat
18	199	50.6	2401	15	AK066680	AK066680 Oryza sat
19	198	50.4	1788	15	AY142017	AY142017 Arabidops
20	198	50.4	2017	15	AY058173	AY058173 Arabidops
21	198	50.4	2057	15	AY056446	AY056446 Arabidops
22	196	49.9	404	6	BD224432	BD224432 Materials
23	196	49.9	404	6	AR216482	AR216482 Sequence
24	196	49.9	404	6	AR432884	AR432884 Sequence
25	196	49.9	1886	15	BT002582	BT002582 Arabidops
26	194	49.9	1921	15	AY062675	AY062675 Arabidops
27	194	49.4	17857	14	AC155503	AC155503 Zea mays
28	191.5	48.7	1926	15	PSCYTP450	Z43263 P.sativum m
29	191	48.6	1077	15	AK220829	AK220829 Arabidops
30	179	45.5	103973	15	AP004028	AP004028 Oryza sat
31	179	45.5	110000	14	CR954215_1	Continuation (2 of
32	179	45.5	110000	15	AP008208_351	Continuation (352
33	179	45.5	145014	15	AP004048	AP004048 Oryza sat
34	177.5	45.2	2269	15	AF459441	AF459441 Skeletone
35	173.5	44.1	197405	15	ATCHRIV40	AL161540 Arabidops
36	173.5	44.1	202860	15	ATFCA2	Z97337 Arabidops
37	159	40.5	378	15	ATH526463	ATH526463 Arabidops
38	152.5	38.8	110000	15	AP008208_039	Continuation (40 o
39	152.5	38.8	170051	15	AP005772	AP005772 Oryza sat
40	149	37.9	119420	15	AC079041	AC079041 Arabidops
41	131	33.3	109093	15	AC148994	AC148994 Medicago
42	131	33.3	120761	15	AC124218	AC124218 Arabidops
43	127	32.3	75683	14	AC153160	AC153160 Glycine m
44	124	31.6	1498	15	AK062235	AK062235 Oryza sat
45	111.5	28.4	86539	15	AP004485	AP004485 Lotus cor

ALIGNMENTS

RESULT 1	AY424805	1620 bp	mRNA	linear	PLN 07-JAN-2004
LOCUS	AY424805	Arabidopsis thaliana chloroplast	carotenoid epsilon-ring		
DEFINITION	hydroxylase (LUT1) mRNA, complete cds; nuclear gene for chloroplast product.				
ACCESSION	AY424805				
VERSION	AY424805.1	GI:40218378			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1620)				
AUTHORS	Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M. and DellaPenna, D.				
TITLE	The Arabidopsis LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid epsilon-ring hydroxylation activity				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (1), 402-407 (2004)				
PUBMED	14709673				
REFERENCE	2 (bases 1 to 1620)				
AUTHORS	Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M. and DellaPenna, D.				
TITLE	Direct Submission				
JOURNAL	Submitted (30-SEP-2003) Biochemistry and Molecular Biology,				

Michigan State University, East Lansing, MI 48824, USA	
FEATURES	Location/Qualifiers
source	1..1620
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosome="3"
	/ecotype="Columbia"
gene	1..1620
	/gene="LUT1"
CDS	1..1620
	/gene="LUT1"
	/codon_start=1
	/product="chloroplast carotenoid epsilon-ring hydroxylase"
	/protein_id="AAR83120.1"
	/db_xref="GI:40218379"
	/translation="MESSLSPSSSYSLFTAKPTLLSPKPKFTFSIRSSIEKPKP KLETNSSQSVSDWTLTLTRLSGKNDSEGIPIANAKLDVADLGGALFLPLY KMNBYGPIYRLAAGPRNFVSDPAIAKHVLRNPYKAGLVAEVSFLLGSGFAIA EGPLTARRAVPSLHRYLSVVERVCKAEKRLVEKLPYAEKGSVNMKAFPSQ MTLDVIGSLNFYDSTLTPSVIEVYTKAEALRSDDLPLPWKIDALCKIVPRQ VKAERAVTLIRFVETDLAKCEIVERGERINDEEYVNDAPDILRFLIASREEVSS VOLRDDLSMLVAGHETGCVLTWTLVLLSKNSALAKAEQVDRVLEGRNPAPFEDIK EUKYITRCINESMRULPHPPVILIRRAQVDPILPGNYKVTGQIMISVNIHRSSFV EKAEFLPERFDIGAINETNTDFKIFPFGSPKRCVGDQFALMEIIVALAVFLQRL NVELVPDQTISMTCGATHTTNGLYMKVSQR"
ORIGIN	
Alignment Scores:	
Pred. No.:	2,24e-42
Score:	393.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	15
	Length: 1620
	Matches: 77
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0
US-10-751-235-1 (1-77) x AY424805 (1-1620)	
QY	1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db	598 TTTCAGCCTTATGCAGACGAGTGTCTGTAATATGGAAGCGAAGTCTCTCAGATG 657
QY	21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db	658 ACATTTGATGTCATTCGGTGTCTCTTTTAACTACAATTCGATTCCTTGACTACTGAT 717
QY	41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db	718 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 777
QY	61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db	778 CTTCTGCCATATTGGAAGATCGATGATTCGTTGTAAGATAGTCCCGAGACAG 828
RESULT 2	
AY091083	1858 bp mRNA linear PLN 18-SEP-2002
LOCUS	Arabidopsis thaliana putative cytochrome P450 (At3g53130) mRNA, partial cds.
DEFINITION	
ACCESSION	AY091083
VERSION	AY091083.1
KEYWORDS	GI:20258841
SOURCE	FLJ CDNA.
ORGANISM	Arabidopsis thaliana (thale cress)
	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 1858)
AUTHORS	Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
	Direct Submission
	Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
	The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
	Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.
	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
FEATURES	Location/Qualifiers
source	1..1858
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosome="3"
	/clone="RAFL09-78-C14 (R19513)"
	/ecotype="Columbia"
	/note="This clone is in a modified pBluescript vector (PLC-1) as a BamHI/XhoI insert."
gene	<1..1858
	/gene="At3g53130"
CDS	<1..1661
	/gene="At3g53130"
	/codon_start=3
	/evidence=experimental
	/product="putative cytochrome P450"
	/protein_id="AA13903.1"
	/db_xref="GI:20258842"
	/translation="Q1FLHGSFEKSSMESSLFPSSSSSYSLFTAKPTLLSPKPKF TFSISSIEKPKPKLETNSSQSVSDWTLTLTRLSGKNDSEGIPIANAKLDVADLGGALFLPLYKMNBYGPIYRLAAGPRNFVSDPAIAKHVLRNPYKAGLVAEVSFLLGSGFAIAEGPLTARRAVPSLHRYLSVVERVCKAEKRLVEKLPYAEKGSVNMKAFPSQMTLDVIGSLNFYDSTLTPSVIEVYTKAEALRSDDLPLPWKIDALCKIVPRQVKAERAVTLIRFVETDLAKCEIVERGERINDEEYVNDAPDILRFLIASREEVSSVOLRDDLSMLVAGHETGCVLTWTLVLLSKNSALAKAEQVDRVLEGRNPAPFEDIK EUKYITRCINESMRULPHPPVILIRRAQVDPILPGNYKVTGQIMISVNIHRSSFV EKAEFLPERFDIGAINETNTDFKIFPFGSPKRCVGDQFALMEIIVALAVFLQRL NVELVPDQTISMTCGATHTTNGLYMKVSQR"
3'UTR	1662..1858
	/gene="At3g53130"
misc_difference	1842
	/gene="At3g53130"
	/note="compared to genomic sequence"
	/replace="t"

ORIGIN

Alignment Scores: 2.61e-42 Length: 1858
Pred. No.: 393.00 Matches: 77
Score: 393.00
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AY091083 (1-1858)

QY 1 LeuGlnProTyrAlaGluLeuSerAlaValAlaMetGluAlaLysPheSerGlnMet 20
639 TTGCAGCCCTATGACAGACGGAAGTCTGTGAATATGGAAGCGAAGTCTCTCAGATG 698
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
699 ACATTGATGTCATTGGTGTCTCTTTTAACTACATTCGATCTTTGACTACTGAT 758
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLeuGluAlaGluLeuArgSerThrAsp 60
759 AGTCCTGTCTATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCTTCTACTGAT 818
QY 61 LeuLeuProTyrTrpLysIleAlaPalaLeuCysLeuValIleValProArgGln 77
819 CTTCTGCCATATTGGAAGATCGATGATTCATTGTGTAAGATAGTCCCGAGACAG 869

RESULT 3

AF367289 3304 bp mRNA linear PLN 24-MAY-2001
LOCUS Arabidopsis thaliana AT3G53140/T4D2_70 mRNA, complete cds.

AF367289 AF367289.1 GI:14194164

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 3304)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,

Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,

Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 3304)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,

FEATURES

source

1. .3304

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAFL09-13-P07 (R12236)"

/ecotype="Columbia"

1. .2037

2038..3117

/note="caffeic acid O-methyltransferase - like protein"

/codon_start=1

/product="AT3G53140/T4D2_70"

/protein_id="AAK56277.1"

/db_xref="GI:14194165"

/translation="MENSSSRNRRLAIMEANMISVPMISNAARVLGIADAIWNG

GANSPLSAEILPLRLHPSHTTIGDPENLQRLMLTSYGVSEHLVGSIRKYSLT

DVGKTLVTDSSGLSYAAVYLQHQALMRANPLVHTAVVETPPTPKVKGAAAYAYO

GKSEEMGLMQKMSGVSPFMKAILDYGKFSVDILVDVGSAGDCILRMILQQPYN

VREGINFPLPEVAKAPNIPGVTHVGGDMFQVPSADAI FMKVWLTWTDSECKQIMK

NCYNALPVGKLIACEAPNIPKETSRSRTRALLEGDI FVMTIYRTKGKHTBEEFIEL

GLSAGFPFRPEYIDYFYTILFQK"

3118..3304

3'UTR

ORIGIN

Alignment Scores:

Pred. No.: 5.03e-42 Length: 3304

Score: 393.00 Matches: 77

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AF367289 (1-3304)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20

602 TTGCAGCCCTATGACAGACGGAAGTCTGTGAATATGGAAGCGAAGTCTCTCAGATG 661

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40

662 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACATTCGATTCCTTGACTACTGAT 721

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60

722 AGTCCTGTCTATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTGCTTCTACTGAT 781

RESULT 4

ATT4D2

LOCUS

DEFINITION

Arabidopsis thaliana DNA chromosome 3, BAC clone T4D2.

ACCESSION

AL132958

VERSION

AL132958.1

GI:6434245

KEYWORDS

Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1

REFERENCE

AUTHORS

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,

Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,

Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (30-MAR-2001)

Salk Institute Genomic Analysis Laboratory

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,

COMMENT

ATT4D2

92611 bp DNA linear

Arabidopsis thaliana

PLN 16-APR-2005

Arabidopsis thaliana

chromosome 3, BAC clone T4D2.

AL132958

GI:6434245

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1

REFERENCE

AUTHORS

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,

Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,

Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,

Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (30-MAR-2001)

Salk Institute Genomic Analysis Laboratory

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,

Chew, R., Kim, C.J., Koesema, E., Meyers, M.C., Tracy, S.B., Banh, J.,

Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

Weichselgartner,M., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F.
and Salanoubat,M.

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 92611)
EU Arabidopsis sequencing project.
Direct Submission

Submitted (21-DEC-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue
Gaston Crenieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr

Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES

source

Location/Qualifiers

1..92611

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/chromosome="3"

/ecotype="Columbia"

2081..2646

/gene="T4D2.10"

Join(2081..2434,2533..2646)

/gene="T4D2.10"

/note="similarity to BETA-GALACTOSIDASE PRECURSOR,

Lycopersicon esculentum, gb:P48980"

/codon_start=1

/product="putative protein"

/protein_id="CAB64211.1"

/db_xref="GI:6630728"

/db_xref="GOA:Q9SCQ3"

/db_xref="InterPro:IPR000922"

/db_xref="UniProt/TREMBL:Q9SCQ3"

/translation="METLQCRHQHVIFILLVLFHSLFLGLASKIDVSDARGIRIDGG
QKRLFTNPGHQEHAACNEEPDGLPRLTRISCNPEGVYITKINFADYGNFTGCHP
RRNGCGARATWRIIVKKNCLGKCHLLVTDEMFPSPKCKGAPMLAVETCTIA"

2081..2434

/gene="T4D2.10"

/number=1

2435..2532

/gene="T4D2.10"

/number=1

2533..2646

/gene="T4D2.10"

/number=2

3341..8521

/gene="T4D2.20"

Join(3341..3364,3812..3934,4065..4305,4394..4956,
5131..6090,6168..6490,6606..6884,6997..7249,7359..7529,
7645..7737,7816..7923,8048..8164,8257..8346,8438..8521)

/gene="T4D2.20"

/note="similarity to KIAA0010 protein, Homo sapiens,

gb:A38919"

/codon_start=1

/product="putative protein"

/protein_id="CAB64212.1"

/db_xref="GI:6630729"

/db_xref="GOA:Q9SCQ2"

/db_xref="InterPro:IPR000048"

/db_xref="InterPro:IPR000569"

/db_xref="UniProt/TREMBL:Q9SCQ2"

/translation="MDLNRRKXVSLRGASSGEISRDALLAKVSOERLRSYARRANAA
SLFIQWRWSYIVRKAALIEIQEWENLLSCHSVTLTKSWSSRVLRPPLFFVRSLSV
HQKIQAREIHCMTCKILLESINNDQYNFCSLAVGTSEDSKTWACQTRMVSIC
SPLLTCTNYSYERIKDVGYNALLRLIVLTPDKSWKIITNFENFSDATKILIIQF
IGSKGYSVAVRYIKTLTKHTDERLVIITSAVLALRPFHVKQPAFVDDNQDFNL
AVEYSVLITLIPRLVCYLSALIRALKHKSILMPSFHTILLKDKILNISSENSE
KOSTWEIIPSGVIGNIISLATVSETDFMDQSPENPFYLVHVIVTLAENLLSQ
VESVGQDILHIDEATSNETKGNYSKISFVEMLRPVCOQHLAKLLASGKEIRVIA
DKDASTSSKSGSETGLGLLDIARLYSCMLRIFCVMPNVLGLPLVNLMSLSPCGYIVSLM

NSLESVLLPENGCTADDASHGSAKTSWNTRSPSEKCLKHLKNDVSNKWNVNLKPSGK
SPGPREHVCTSQDPGSGQVNESTNDVMDVETLGRGPGVIGISKEYSCLLHLFCATYAHLL
LVLDLDDIQYEKQPFMLEKQORIASMLNTLVYGLLRTGPGESRQMLDSAIRCLHLL
YERDCRPFPCASALMLSPGRTSREPIAFAARTHEVLPTSDVLTTPSGSVITITPHVF
PFEEHVHREFISKORASRMAGEVDAPGARSIEIVVRGHVVVEDGFOQLNSIGSRL
KSSIHVSFVNESGLPEAGLDYGGUSKEFLDITTKAAPATEYGLSQPTSRLLVPSPP
SARLLENGIQMIEFLGRIVGKALYEGILLDYFSHVFIOKLGRYSFIDELSGLDPEL
YRLNMYKHYQDGLKEILCLDFTVTEEFCKGMSITELKPGGKDTSVTNENKQYTHAMA
DYKLNROIVPNSAFYRGLTDLISPAWLKLFNAHEFNOLLSGGNHDI DVDDLRRNTKY
TGYSDSRIKIFWEVMKGFEPSERCILLKFKVTCSCRAPLLGPKYLOPTFIHKVSC
DTSUWAAIGQDQVERLPFSATCYNTLKLPTFKRASTWREKLLYALTSNAGFELS"
3341..3364
/gene="T4D2.20"
/number=1
3365..3811
/gene="T4D2.20"
/number=1
3812..3934
/gene="T4D2.20"
/number=2
3935..4064
/gene="T4D2.20"
/number=2
4065..4305
/gene="T4D2.20"
/number=3
4306..4393
/gene="T4D2.20"
/number=3
4394..4956
/gene="T4D2.20"
/number=4
4957..5130
/gene="T4D2.20"
/number=4
5131..6090
/gene="T4D2.20"
/number=5
6091..6167
/gene="T4D2.20"
/number=5
6168..6490
/gene="T4D2.20"
/number=6
6491..6605
/gene="T4D2.20"
/number=6
6606..6884
/gene="T4D2.20"
/number=7
6885..6996
/gene="T4D2.20"
/number=7
6997..7249
/gene="T4D2.20"
/number=8
7250..7358
/gene="T4D2.20"
/number=8
7359..7529
/gene="T4D2.20"
/number=9
7530..7644
/gene="T4D2.20"
/number=9
7645..7737
/gene="T4D2.20"
/number=10
7738..7815
/gene="T4D2.20"
/number=10
7816..7923
/gene="T4D2.20"

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

```

/number=11
7324. .8047
/genes="T4D2.20"
/number=11
8048. .8164
/genes="T4D2.20"
/number=12
8165. .8256
/genes="T4D2.20"
/number=12
8257. .8346
/genes="T4D2.20"
/number=13
8347. .8437
/genes="T4D2.20"
/number=13
8438. .8521
/genes="T4D2.20"
/number=14
9057. .10601
/genes="T4D2.30"
complement(join(9057. .9151,9403. .9658,9810. .10043,
10124. .10254,10370. .10601))
/genes="T4D2.30"
complement(join(9057. .9151,9403. .9658,9810. .10043,
10124. .10254,10370. .10601))
/genes="T4D2.30"
/notes="strong similarity to proline-rich protein, Raphanus
sativa, PIR:S16748"
/codon_start=1
/product="putative protein"
/protein_id="CAB64213.1"
/db_xref="GI:6630730"
/db_xref="GOA:Q9SCQ1"
/db_xref="InterPro:IPR001087"
/db_xref="UniProt/TREMBL:Q9SCQ1"
/translation="MRVSGFRVLLVSCFFCKSKGVNPALIMFGDSIVDVGNNNLL
SIVKSNFLPYGRDFIDQPTGRFCNGKLAIDFSAEYLGFSYPAPLFSREASNNELI
GANFASNSGYDAYSVPFGSISLTQLSYRAYQNVYRMIGRNARILFSGRHIL
SAGSDFIQNYINFLNLPDQFADILLRSFEFQNLNLYELGARIGVTSLPMPG
CLPAATILFGAGNCKSVRLNNDALFMENTKLENTTLLMNRHSLRLVAENVYQPLD
IITPTDNGYVPWDFPHFTEAVNELLAQQLGQGISLIN"

Alignment Scores:
Pred. No.: 4,246-34 Length: 92611
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: 15 Gaps: 1

US-10-751-235-1 (1-77) x ATT4D2 (1-92611)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
DB 17626 TTGCAGCCTTATGCAGAGACCAAGTCTGTGAATATGGAAGCAAGTTCTCTCAGATG 17685

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
DB 17686 ACACCTTGATGCTGGTGTCTCTTTTAACTACAAATTCGATCTTTGACTACTGAT 17745

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 17746 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTGCTTCTACTGAT 17805

QY 61 LeuLeuProTyrTrpLysIle----- 67
DB 17806 CTTCTGCCATATTGGAAGGC-AAGTTCTCTGTGTTTTTCTGTGTTTGTGTTGTAATGTGTG 17864

QY 68 -----AspAlaLeu 70
DB 17865 GAACAAATTGGATCTTTGTTAATTGAGAGGGTTTGGTTGTTTTCAGATGATGCAATG 17924
```

```

QY 71 CysLysIleValProArgGln 77
DB 17925 TGTAAGATAGTCCCGAGACAG 17945

RESULT 5
LOCUS BT012891
DEFINITION Lycopersicon esculentum clone 113997P, mRNA linear PLN 11-MAY-2004
ACCESSION BT012891
VERSION BT012891.1 GI:47104306
KEYWORDS FLI CDNA.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1967)
REFERENCE 1 (bases 1 to 1967)
AUTHORS Kirnase, E.P., Wang, W. and Vazeille, A.
DIRECT SUBMISSION
TITLE Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
JOURNAL Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1. 1967
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="113997P"
/tissue_type="trichome"
/notes="HTDI88"

ORIGIN
Alignment Scores:
Pred. No.: 8,546-35 Length: 1967
Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x BT012891 (1-1967)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
DB 604 CTTTACCTGATGAGCAATTTCTGGCTCTGCAGTGAATATGAGGCAAGTTTCTCACTA 663

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
DB 664 ACACCTTGATGCTGGTGTCTCTTTTAACTACAAATTCGATCTTTGACTACTGAT 723

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 724 AGTCAGTATTATGATGCAAGTTTACACTGCACTTAAAGAAGCAAGTCCGTTCACTGAT 783

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 784 TTGTTGCCATATTGCGAGATCAAGCTTTATGTAAAGTTTCATCCACGACAA 834

RESULT 6
LOCUS AK065689
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013039G17, full
insert sequence.
ACCESSION AK065689
VERSION AK065689.1 GI:32975707
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team;
```

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nakikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayaahizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

12869764

2 (bases 1 to 1876)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayaahizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, K., Satoh, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sogabe, Y., Sugano, S., Shinagawa, A., Shiraki, T., Shishiki, T., Shishiki, T., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,

Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayaahizaki, Y.

Location/Qualifiers

source

1. 1876

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J013039G17"

ORIGIN

Alignment Scores:

Pred. No.: 15e-34 Length: 1876

Score: 335.00 Matches: 66

Percent Similarity: 90.91% Conservative: 4

Best Local Similarity: 85.71% Mismatches: 7

Query Match: 85.24% Indels: 0

DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AK065689 (1-1876)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20

Db 726 CTTGAGACATCTGCTTTAAAGTGCGAAACCTGTAAATATGGAAGCAAGGTCTCTCAATG 785

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40

Db 786 ACTTAGAATGATGGTTGCTTGTTCCTCAATACATTTGATTCCCTCACATCAGAT 845

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60

Db 846 AGCCTGTATGTGCTGTTTACACTGCACCTCAAGGAAGCAGAACTTCGTTCTACAGAT 905

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77

Db 906 CTTTACCATACATGGAAGATTGATTGCTGTGCAAGATTGTTCTTAGACAA 956

RESULT 7

AP008216_208

WPCOMMENT

Sequence split into 227 fragments LOCUS AP008216 Accession AP008216

Fragment Name	Begin	End
AP008216_000	1	110000
AP008216_001	100001	210000
AP008216_002	200001	310000
AP008216_003	300001	410000
AP008216_004	400001	510000
AP008216_005	500001	610000
AP008216_006	600001	710000
AP008216_007	700001	810000
AP008216_008	800001	910000
AP008216_009	900001	1010000
AP008216_010	1000001	1110000
AP008216_011	1100001	1210000
AP008216_012	1200001	1310000
AP008216_013	1300001	1410000
AP008216_014	1400001	1510000
AP008216_015	1500001	1610000
AP008216_016	1600001	1710000
AP008216_017	1700001	1810000
AP008216_018	1800001	1910000
AP008216_019	1900001	2010000
AP008216_020	2000001	2110000
AP008216_021	2100001	2210000
AP008216_022	2200001	2310000
AP008216_023	2300001	2410000
AP008216_024	2400001	2510000
AP008216_025	2500001	2610000
AP008216_026	2600001	2710000
AP008216_027	2700001	2810000
AP008216_028	2800001	2910000
AP008216_029	2900001	3010000
AP008216_030	3000001	3110000

AP008216_031 3100001 3210000
AP008216_032 3200001 3310000
AP008216_033 3300001 3410000
AP008216_034 3400001 3510000
AP008216_035 3500001 3610000
AP008216_036 3600001 3710000
AP008216_037 3700001 3810000
AP008216_038 3800001 3910000
AP008216_039 3900001 4010000
AP008216_040 4000001 4110000
AP008216_041 4100001 4210000
AP008216_042 4200001 4310000
AP008216_043 4300001 4410000
AP008216_044 4400001 4510000
AP008216_045 4500001 4610000
AP008216_046 4600001 4710000
AP008216_047 4700001 4810000
AP008216_048 4800001 4910000
AP008216_049 4900001 5010000
AP008216_050 5000001 5110000
AP008216_051 5100001 5210000
AP008216_052 5200001 5310000
AP008216_053 5300001 5410000
AP008216_054 5400001 5510000
AP008216_055 5500001 5610000
AP008216_056 5600001 5710000
AP008216_057 5700001 5810000
AP008216_058 5800001 5910000
AP008216_059 5900001 6010000
AP008216_060 6000001 6110000
AP008216_061 6100001 6210000
AP008216_062 6200001 6310000
AP008216_063 6300001 6410000
AP008216_064 6400001 6510000
AP008216_065 6500001 6610000
AP008216_066 6600001 6710000
AP008216_067 6700001 6810000
AP008216_068 6800001 6910000
AP008216_069 6900001 7010000
AP008216_070 7000001 7110000
AP008216_071 7100001 7210000
AP008216_072 7200001 7310000
AP008216_073 7300001 7410000
AP008216_074 7400001 7510000
AP008216_075 7500001 7610000
AP008216_076 7600001 7710000
AP008216_077 7700001 7810000
AP008216_078 7800001 7910000
AP008216_079 7900001 8010000
AP008216_080 8000001 8110000
AP008216_081 8100001 8210000
AP008216_082 8200001 8310000
AP008216_083 8300001 8410000
AP008216_084 8400001 8510000
AP008216_085 8500001 8610000
AP008216_086 8600001 8710000
AP008216_087 8700001 8810000
AP008216_088 8800001 8910000
AP008216_089 8900001 9010000
AP008216_090 9000001 9110000
AP008216_091 9100001 9210000
AP008216_092 9200001 9310000
AP008216_093 9300001 9410000
AP008216_094 9400001 9510000
AP008216_095 9500001 9610000
AP008216_096 9600001 9710000
AP008216_097 9700001 9810000
AP008216_098 9800001 9910000
AP008216_099 9900001 10010000
AP008216_100 10000001 10110000
AP008216_101 10100001 10210000
AP008216_102 10200001 10310000
AP008216_103 10300001 10410000

AP008216_104 10400001 10510000
AP008216_105 10500001 10610000
AP008216_106 10600001 10710000
AP008216_107 10700001 10810000
AP008216_108 10800001 10910000
AP008216_109 10900001 11010000
AP008216_110 11000001 11110000
AP008216_111 11100001 11210000
AP008216_112 11200001 11310000
AP008216_113 11300001 11410000
AP008216_114 11400001 11510000
AP008216_115 11500001 11610000
AP008216_116 11600001 11710000
AP008216_117 11700001 11810000
AP008216_118 11800001 11910000
AP008216_119 11900001 12010000
AP008216_120 12000001 12110000
AP008216_121 12100001 12210000
AP008216_122 12200001 12310000
AP008216_123 12300001 12410000
AP008216_124 12400001 12510000
AP008216_125 12500001 12610000
AP008216_126 12600001 12710000
AP008216_127 12700001 12810000
AP008216_128 12800001 12910000
AP008216_129 12900001 13010000
AP008216_130 13000001 13110000
AP008216_131 13100001 13210000
AP008216_132 13200001 13310000
AP008216_133 13300001 13410000
AP008216_134 13400001 13510000
AP008216_135 13500001 13610000
AP008216_136 13600001 13710000
AP008216_137 13700001 13810000
AP008216_138 13800001 13910000
AP008216_139 13900001 14010000
AP008216_140 14000001 14110000
AP008216_141 14100001 14210000
AP008216_142 14200001 14310000
AP008216_143 14300001 14410000
AP008216_144 14400001 14510000
AP008216_145 14500001 14610000
AP008216_146 14600001 14710000
AP008216_147 14700001 14810000
AP008216_148 14800001 14910000
AP008216_149 14900001 15010000
AP008216_150 15000001 15110000
AP008216_151 15100001 15210000
AP008216_152 15200001 15310000
AP008216_153 15300001 15410000
AP008216_154 15400001 15510000
AP008216_155 15500001 15610000
AP008216_156 15600001 15710000
AP008216_157 15700001 15810000
AP008216_158 15800001 15910000
AP008216_159 15900001 16010000
AP008216_160 16000001 16110000
AP008216_161 16100001 16210000
AP008216_162 16200001 16310000
AP008216_163 16300001 16410000
AP008216_164 16400001 16510000
AP008216_165 16500001 16610000
AP008216_166 16600001 16710000
AP008216_167 16700001 16810000
AP008216_168 16800001 16910000
AP008216_169 16900001 17010000
AP008216_170 17000001 17110000
AP008216_171 17100001 17210000
AP008216_172 17200001 17310000
AP008216_173 17300001 17410000
AP008216_174 17400001 17510000
AP008216_175 17500001 17610000
AP008216_176 17600001 17710000


```

/chromosome="10"
/map="near C809"
/clone="OSUNBa0001014"
/notes="Japonica cultivar-group"
660. .1022
/gene="OSJNBa0001014.22"
/notes="predicted by genscan"
4660. .1022
/gene="OSJNBa0001014.22"
660. .1022
/gene="OSJNBa0001014.22"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK20056.1"
/db_xref="GI:13357259"
/translation="MTPNPMQCGFPFRSTRHRQPLRHGGGGVTVVSRGGGGGGA
RVKIVGDELDRIAAGVTRQCAVSPPELLQQARATEVAAPMPMRPESEGGALV
AGRRGEWOPALDGIPEEA"
repeat_region 766. 786
/rpt_family="(CGG)n"
repeat_region 3343. .3469
/rpt_family="putative Stowaway type MITE element"
repeat_region complement(3354. .3480)
/rpt_family="putative Stowaway type MITE element"
repeat_region complement(3808. .3839)
/rpt_family="AT-rich"
repeat_region complement(3937. .3963)
/rpt_family="AT-rich"
repeat_region complement(4836. .4871)
/rpt_family="(TA)n"
misc_feature 4845. .4870
/notes="Overlap BAC OSUNBa0063E23 (AC113947) has 17 AT
repeats in this region."
repeat_region 5433. .5567
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
repeat_region 5441. .10390
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
repeat_region 5499. .5636
/rpt_family="putative Tourist type MITE element"
repeat_region complement(6712. .6743)
/rpt_family="AT-rich"
repeat_region complement(9156. .9226)
/rpt_family="retroelement from gi|5295936|AB026295 Oryza
sativa (160208 to 165872) 5665 nt"
repeat_region 10278. .10452
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
repeat_region complement(10291. .10434)
/rpt_family="Tourist_O11 MITE element from gb:U72728 Oryza
longistaminata receptor-like kinase protein (Xa21), family
member F, pseudogene sequence (1080 to 1236) 157 nt"
repeat_region 10300. .10397
/rpt_family="putative Tourist type MITE element"
repeat_region 10324. .10474
/rpt_family="putative Tourist type MITE element"
repeat_region complement(12414. .12457)
gene 15036. .17810
/notes="OSJNBa0001014.25"
/mRNA join(<15036. .15183,15255. .15433,16442. .16525,17234. .17401,
17688. .>17810)
/genes="OSJNBa0001014.25"
/CDS join(15036. .15183,15255. .15433,16442. .16525,17234. .17401,
17688. .17810)
/genes="OSJNBa0001014.25"
/codon_start=1
/product="hypothetical protein"

```

```

/protein_id="AAK20044.1"
/db_xref="GI:13357247"
/translation="MDASPSSSSSPVAPEVVQWAGADGFGSAAIGFACRYLCAVAT
CVFAAVGVVAGTGAICLATESGLRGAGTGAISGAVFSIEVAESSRDLWHSSDSA
VWCLVYMDYIISSLSGRLEKRGKVPQVQSAVQSOISAISSPASTSDLFETGGTKGL
PADTLRLPAIKINGDNAVDSAGBPICCSVCLQDFRVGEMARBLPSCHRVFHVPCIDC
WLVRHGSCPLCRRDI"
repeat_region complement(20477. .20596)
/rpt_family="Gaifjin_Oe1 MITE element from gb:X52623 Rice
4-CL gene for 4-coumarate-CoA ligase (1888 to 2027) 140
nt"
repeat_region 22465. .22516
/rpt_family="(GAAA)n"
repeat_region complement(23182. .23375)
/rpt_family="Rice RS081X19 insert DNA with repeat units
R1-R6 gi|20362|emb|X5540.1|QSRSR1R6"
repeat_region complement(23216. .23353)
/rpt_family="DNA fragment containing the junction of AA
genome-specific tandem repeat sequence TrsA
(telomere-associated) gi|804936|dbj|D16452.1|RICTRSA1 Rice
(O.sativa), TrsA1 locus"
repeat_region complement(23490. .23515)
/rpt_family="AT-rich"
repeat_region complement(23739. .23847)
/rpt_family="(CGG)n"
repeat_region 23861. .24031
/rpt_family="(CCGCG)n"
repeat_region complement(23942. .24112)
/rpt_family="(CGG)n"
repeat_region complement(24108. .24214)
/rpt_family="(AGGGGG)n"
repeat_region complement(24234. .24404)
/rpt_family="(CGG)n"
repeat_region 24521. .24618
/rpt_family="(GAAA)n"
repeat_region complement(24651. .24680)
/rpt_family="AT-rich"
repeat_region complement(27603. .27652)
/rpt_family="GC-rich"
repeat_region 27711. .27833
/rpt_family="(CCGCG)n"
repeat_region complement(28641. .29130)

Alignment Scores:
Pred. No.: 4.99e-28 Length: 178022
Score: 303.00 Matches: 66
Percent Similarity: 57.85% Conservative: 4
Best Local Similarity: 54.55% Mismatches: 7
Query Match: 77.10% Indels: 44
DB: 15 Gaps: 1

US-10-751-235-1 (1-77) x AC025783 (1-178022)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 113284 CTTGAGACATCTGCTTTAAGTGCGCAACCTGTAATATGGAAGCACTCTCTCAATG 113343
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 113344 ACTTAGATGATGTTGTTGCTCTTGTTCATTTACATTTTGTTCCTTCATCAGAT 113403
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 113404 AGCCTGTATTGATGCTGTTTACTGCTCACTCAAGGAAGCAGAACTCGTCTCAGAT 113463
QY 61 LeuLeuProTyrTrpLys-----
Db 113464 CTTTACCATATCGAAGGTAATGTTCTTTGCCAAGTTATTTTCTAAATAATCAAATT 113523
QY 66 -----
Db 113524 ATGGTTGCGCCTATCTAACTTCTAGTGACGAAATTCGTGACTGATGTCGCTGACTGATG 113583

```

Qy 67 -----llesapalateuCyLysllevalproArg 76
Db 113584 TCTGATGAGCCCTGTTTGTCTTGATAGATTGCTGTGCAAGATTGTTCTCTAGA 113643
Qy 77 Gln 77
Db 113644 CAA 113646
RESULT 9 299735 bp DNA linear PLN 06-JUN-2003
LOCUS AE017117
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 10, section 71 of
77 of the complete sequence.
ACCESSION AE017117 AE016959
VERSION AE017117.1 GI:31433358
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 299735)
The Rice Chromosome 10 Sequencing Consortium
In-depth view of structure, activity, and evolution of rice
chromosome 10
Science 300, 1566-1569 (2003)
2 (bases 1 to 299735)
Buell,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q.
Direct Submission
Submitted (05-MAY-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT
This is the pseudomolecule for rice chromosome 10, which was
constructed by resolving discrepancies between overlapping BACs,
trimming the overlap regions, and linking the unique sequences to
form a contiguous sequence. Genes in individual BAC clone were
identified by a combination of several methods: Gene prediction
programs, searches of the complete sequence against a peptide
database and EST databases. Genes with similarity to other proteins
are named after the database hits. Genes without significant
peptide similarity but with EST similarity are named as unknown
proteins. Genes without protein or EST similarity, that are
predicted by more than two gene prediction programs over most of
their length are annotated as hypothetical proteins. Genes
encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES
source
1..299735 Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
misc_feature
<1..133260
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSUNBa0001014
(GB:AC025783)."
<431..>3685
/locus_tag="OSUNBa0001014.2"
/notes="EST D48843, BE230623 from this gene"
join(431..633,699..781,864..905,1063..1104,1442..1525,
1702..1789,1913..2010,3644..3685)
/locus_tag="OSUNBa0001014.2"
join(471..633,699..781,864..905,1063..1104,1442..1525,
1702..1789,1913..2010,3644..3685)
/locus_tag="OSUNBa0001014.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP54878.1"
/db_xref="GI:31433359"
/translation="MAMAATPVLLLRSSATILLPRSSPRASACVLAPQRRRGRWR
RLRSSPPPEGVPGELMESEDSKFPVLAEDPMYGPALLIGFEEGTNKVQEFRLR
ELDGEFLKVIHCTEEMTKQLWDAMTEQPSIEAVKIASMPRICIFSGLTGEMMWF

gene
INSPYPTGLEPFAAFAALVPNSADKILGEVIEEINGDHEMLLLVAGDFSRSCYC"
<6954..>9818
/locus_tag="OSUNBa0001014.3"
/note="similar to C 4 sterol methyl oxidase GB:AAAC49139
GI:1161339 (Saccharomyces cerevisiae); EST D15076, C97849
from this gene"
join(6954..7365,7735..8057,9074..9221,9316..9348,
9516..9818)
/locus_tag="OSUNBa0001014.3"
join(7036..7365,7735..8057,9074..9221,9316..9348,
9516..9587)
/locus_tag="OSUNBa0001014.3"
/codon_start=1
/product="putative C-4 sterol methyl oxidase"
/protein_id="AAP54879.1"
/db_xref="GI:31433360"
/translation="MLPYATAAEAEALGRAMTAESLWFRYSAGIDPYVLFWNHILF
LFVFTLAPLPALELRAPAAVGFPKLQKVRLSREFFRCYRDVMDKFLFVIGPLQ
LVSPTVMVGIIHTGLPLPSIGENAAQLVFLVEDLYLNYIHLLHGEWGEYKIHVS
HHEFTAPIGFAAPVAHAEVLILGIPSGFVGALAPGHMITFELWIVLQMEAIETHSG
FDFFNLTKYIPFYGAEYHDYHVGVRQSQSNFASVPTYCDLYLGTDKGYRYHKAYQ
AKMALGQTEGEKADSNGLSYAKLD"
complement(<10045..>13873)
/locus_tag="OSUNBa0001014.4"
/note="contains pfam profile PF00098 (zinc finger, CCHC
class); EST AU161993, AU057223, AU093006, AU057224 from
this gene"
complement(join(<10045..10997,11698..11793,12830..13057,
13670..13873))
/locus_tag="OSUNBa0001014.4"
complement(join(10454..10997,11698..11793,12830..12933))
/locus_tag="OSUNBa0001014.4"
/codon_start=1
/product="putative zinc finger protein"
/protein_id="AAP54880.1"
/db_xref="GI:31433361"
/translation="MSSRSPPPKDKRIRITRTSYRDAPYRDRSRGPRFPNDLCNCC
KRGPHARDCPNMLCHACGLPGHIAECSKDLCKWCKEPGHMANSFNEGICRNCG
KSGHTARECSAPNMLGMRCLSCNYKFGHLAAECTNEKACNNCKSGHLARNCPEP
VCNLVNSGHLARECPKSDAINEERGPPPPFRRGGYSDVVCACRACQVGHMSRDCMAGAFM
ICHNCGGGRGHAYECPSGRLMDRFPFRRF"
<14813..>17815
/locus_tag="OSUNBa0001014.12"
/note="predicted by fgenesh"
join(<14813..15352,16307..16418,16786..16950,
17391..>17815)
/locus_tag="OSUNBa0001014.12"
join(14813..15352,16307..16418,16786..16950,17391..17815)
/locus_tag="OSUNBa0001014.12"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54881.1"
/db_xref="GI:31433362"
/translation="MALVGCNSGEAAAPAAWTTPYCTVVAADMDSFCVLSCPRCERA
LPDHADAACTGRCGGPGVPARVYRLASVATHDRVVPVLFPAARVLYGCPDAE
LARFAAHAGARAAEEALEGCVKAVMAFAKGAERFRAVSVPVLRDGFPLIDTL
RELYCTADPTPATSPPLEFTVFLVYVVKRKKVYVTLDSRRKIRTESIDLAQLVEA
TMDFSKQITLDFDVTIERLRNEFSNVSEVPALIALADPTCTVQHPPLNRSIQHER
NLPLNSQTLTSCVRCGLLELRAACKRHEADQSHSDAPRTCTVQHPPLNRSIQHER
LCGALNMDIDPEVLVTKMTQCEQRTMNCISMLDQFSFAFNKHVLTITSCMCIVWFL
LYYWRKSYVPEVYHDTSSAIN"
<18650..>20461
/locus_tag="OSUNBa0001014.5"
/note="similar to cellulase GB:CAA48324 GI:311835
(Tropaeolum majus); EST C24984 from this gene"
join(18650..18935,19046..19146,19237..19454,20098..>20461)
/locus_tag="OSUNBa0001014.5"
join(18698..18935,19046..19146,19237..19454,20098..>20461)
/locus_tag="OSUNBa0001014.5"
/codon_start=1
/product="putative cellulase"
/protein_id="AAP54882.1"
/db_xref="GI:31433363"

----- Project informaton -----	
Web site: http://www.tigr.org/tdb/tgi/maize/	
Contact: maize@tigr.org	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 31 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
1	1702: contig of 1702 bp in length
*	1703
*	1802: gap of unknown length
*	1803
*	4662: contig of 2860 bp in length
*	4663
*	4762: gap of unknown length
*	4763
*	11547: contig of 6785 bp in length
*	11548
*	11647: gap of unknown length
*	15558: contig of 3911 bp in length
*	15559
*	15658: gap of unknown length
*	15659
*	18072: contig of 2414 bp in length
*	18073
*	18172: gap of unknown length
*	18173
*	22715: contig of 4543 bp in length
*	22716
*	22815: gap of unknown length
*	35446: contig of 12631 bp in length
*	35447
*	35546: gap of unknown length
*	35547
*	59875: gap of unknown length
*	59876
*	61761: contig of 1786 bp in length
*	61762
*	61861: gap of unknown length
*	64113: contig of 2252 bp in length
*	64114
*	64213: gap of unknown length
*	68252: contig of 4039 bp in length
*	68253
*	68352: gap of unknown length
*	71635: contig of 3283 bp in length
*	71636
*	71735: gap of unknown length
*	74351: contig of 2616 bp in length
*	74352
*	74451: gap of unknown length
*	74352
*	78944: contig of 4493 bp in length
*	78945
*	79044: gap of unknown length
*	78945
*	82988: contig of 3944 bp in length
*	82989
*	83088: gap of unknown length
*	84257: contig of 1169 bp in length
*	84258
*	84357: gap of unknown length
*	84358
*	88368: contig of 4011 bp in length
*	88369
*	88468: gap of unknown length
*	88469
*	103420: contig of 14952 bp in length
*	103421
*	103520: gap of unknown length
*	103521
*	111900: contig of 8380 bp in length
*	111901
*	112000: gap of unknown length
*	112001
*	139514: contig of 27514 bp in length
*	139515
*	139614: gap of unknown length
*	141412: contig of 1798 bp in length
*	141413
*	141512: gap of unknown length
*	141513
*	146353: contig of 4841 bp in length
*	146354
*	146454: gap of unknown length
*	150398: contig of 3945 bp in length
*	150399
*	150498: gap of unknown length
*	152460: contig of 1962 bp in length
*	152461
*	152560: gap of unknown length
*	152561
*	153779: contig of 1219 bp in length
*	153780
*	153879: gap of unknown length
*	167812: contig of 13933 bp in length
*	167813
*	167912: gap of unknown length
*	167913
*	170602: contig of 2690 bp in length
*	170603
*	170702: gap of unknown length
*	170703
*	198608: contig of 27906 bp in length
*	170703
*	198609
*	198708: gap of unknown length
*	198709
*	201017: contig of 2309 bp in length
*	201018
*	201117: gap of unknown length
*	201118
*	204279: contig of 3162 bp in length
*	204280
*	204379: gap of unknown length
*	205408: contig of 1029 bp in length.

FEATURES	Location/Qualifiers
source	1. .205408
	/organism="Zea mays"
	/mol_type="genomic DNA"
	/strain="B73"
	/db_xref="taxon:4577"
	/clone="ZMMBBC0171N08"
	1703. .1802
gap	/estimated_length=unknown
gap	4663. .4762
gap	/estimated_length=unknown
gap	11548. .11647
gap	/estimated_length=unknown
gap	15559. .15658
gap	/estimated_length=unknown
gap	18073. .18172
gap	/estimated_length=unknown
gap	22716. .22815
gap	/estimated_length=unknown
gap	35447. .35546
gap	/estimated_length=unknown
gap	59876. .59975
gap	/estimated_length=unknown
gap	61762. .61861
gap	/estimated_length=unknown
gap	64114. .64213
gap	/estimated_length=unknown
gap	68253. .68352
gap	/estimated_length=unknown
gap	71636. .71735
gap	/estimated_length=unknown
gap	74352. .74451
gap	/estimated_length=unknown
gap	78945. .79044
gap	/estimated_length=unknown
gap	82989. .83088
gap	/estimated_length=unknown
gap	84258. .84357
gap	/estimated_length=unknown
gap	88369. .88468
gap	/estimated_length=unknown
gap	103421. .103520
gap	/estimated_length=unknown
gap	111901. .112000
gap	/estimated_length=unknown
gap	139515. .139614
gap	/estimated_length=unknown
gap	141413. .141512
gap	/estimated_length=unknown
gap	146354. .146453
gap	/estimated_length=unknown
gap	150399. .150498
gap	/estimated_length=unknown
gap	152461. .152560
gap	/estimated_length=unknown
gap	153780. .153879
gap	/estimated_length=unknown
gap	167813. .167912
gap	/estimated_length=unknown
gap	170603. .170702
gap	/estimated_length=unknown
gap	198609. .198708
gap	/estimated_length=unknown
gap	201018. .201117
gap	/estimated_length=unknown
gap	204280. .204379
gap	/estimated_length=unknown
ORIGIN	
Alignment Scores:	
Pred. No.:	8.48e-18
Score:	227.00
Percent Similarity:	80.88%
	Length: 205408
	Matches: 49
	Conservative: 6

Seoul 151-742, Republic of Korea
Location/Qualifiers
1. .2038
/organism="Ginkgo biloba"
/mol_type="mRNA"
/db_xref="taxon:3311"
5'UTR
CDS
11..1771
/codon_start=1
/product="putative 97B2-like cytochrome P450"
/protein_id="AA28222.1"
/db_xref="GI:47498772"
/translation="MITVRSIPCSSEKDTVPQIGRKVLTVENVVSCGVSSSFRLPGS
KPPRCESSTERRAKSNRTLLDNASNFLNLLSGQLGTMEIAGVNSDLFGKELFF
SYDPIEHSVTKLAFGPKSFVVVSDPIVVVHILKFNAPAYDKGVLADILPEIMKKG
LIPADIGTWKPKRAIVPGFHAFFLEAMVKVFDCSERTINLKQLSLLAAEADKTMHI
DVMFAEFNLALDIIIGSVFNDFGVSFVRESPIKAVYGTILFEAHRSTFYIPYWK
PLARMLVPQRKFHEDLKIINECLDSLIQAKETQEDDI EALQGRDYSKVQDASLLR
FLVDKGVVDVNGQLRDDLMTWLIAGHETTAVALTWALFLAONTDKLVKAQAEIDTI
LQQRKTFEDIKRLQYIRLIIAALRLYQPPELLIRRLRQDTIFGGYRGDKDGHILP
KGTDIPISVYNLHRSFYWDQPHFPRPERFLIPKONKGLGEGWAGFDYRSQALYPNE
VLADPAFLPFGGPRKCVGDQFALMESTVGLAMLLQKFEVELKDPPESLRIVTGATIH
TQGLWCKIKRRKHYSESLQ"

ORIGIN

Alignment Scores:
Pred. No.: 1.35e-16 Length: 2038
Score: 201.00 Matches: 40
Percent Similarity: 76.12% Conservative: 11
Best Local Similarity: 59.70% Mismatches: 16
Query Match: 51.15% Indels: 0
DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AY601887 (1-2038)

Qy 11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe 30
Db 668 GTGACATGAGGACGAATTTCAATCTAGCTCTTGACATTATTGGCTAAGTGCTTC 727
Qy 31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla 50
Db 728 AATTAGGACTTTGGATCTGTACTAGAGAAATCACCAGTAATTAAGGCAGTCTATGGTACG 787
Qy 51 LeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTrpLysIleAspAlaLeu 70
Db 788 TTGTTTGAAGCTGAGCAGACATCACTTTTACATTCCATCTACTGGAATTTCTCTTGCA 847
Qy 71 CysLysIleValProArgGln 77
Db 848 AGATGGCTAGTTCCTCGCAA 868

Search completed: January 15, 2006, 14:22:26
Job time : 3726 secs

This Page Blank (uspto)

GenCore version 5.1.1.6

Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 20:19:34 ; Search time 12340 Seconds
(without alignments)
11364.084 Million cell updates/sec

Title: US-10-751-235-5

Perfect score: 2467

Sequence: 1 atgaggtcttcactctttc.....tgaaggtgagccaaaggtaa 2467

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pri.*

9: gb_ro.*

10: gb_sts.*

11: gb_sv.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2467	100.0	92611	15	ATT4D2
2	803	32.5	1620	15	AY424805
3	803	32.5	1858	15	AY091083
4	803	32.5	3304	15	AF367289
5	376.4	15.3	378	15	AJ526463
6	268.6	10.9	1077	15	AK220829
7	252	10.2	1967	15	BT012891
8	239.6	9.7	110000	15	AF008216
9	239.6	9.7	178022	15	AC025783
10	239.6	9.7	299735	15	AE017117
11	183	7.4	1876	15	AK065689
12	146.8	6.0	205408	14	AC155564
13	91	3.7	1788	15	AY142017
14	91	3.7	2017	15	AY058173
15	91	3.7	2057	15	AY056446
16	83.4	3.4	4217	15	AK068163
17	75.2	3.0	178657	14	AC155503
18	74.8	3.0	110000	14	CR954209_3

19	73.2	3.0	1831	15	AF022457
20	72.8	3.0	1926	15	PSCYTP450
21	72	2.9	1886	15	BT002582
22	72	2.9	1921	15	AY062675
23	71.8	2.9	110000	14	CR954213_4
24	71.4	2.9	110000	14	CR954201_08
25	71.2	2.9	625	15	AK221212
26	70.4	2.9	119420	15	AC079041
27	67	2.7	2093	15	AK100596
28	67	2.7	2401	15	AK066680
29	66.4	2.7	181167	14	AC155486
30	66.2	2.7	2038	15	AY601887
31	65.2	2.6	103973	15	AP004028
32	65.2	2.6	110000	15	AP008208_351
33	65.2	2.6	145014	15	AP004048
34	59.4	2.4	197405	15	ATCHRIV40
35	59.4	2.4	202860	15	ATFCA2
36	59	2.4	109093	15	AC148994
37	59	2.4	120761	15	AC124218
38	58.2	2.4	110000	14	CR954215_1
39	57.8	2.3	404	6	BD224432
40	57.8	2.3	404	6	AR216482
41	57.8	2.3	404	6	AR432884
42	56.6	2.3	2269	15	AF459441
43	55.2	2.2	207718	14	AC155487
44	53.8	2.2	86539	15	AP004485
45	53	2.1	110000	1	CP000001_30

ALIGNMENTS

RESULT 1

ATT4D2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 92611
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/ecotype="Columbia"
2081. 2646
/gene="T4D2.10"
join(2081. 2434,2533. 2646)
/gene="T4D2.10"

gene

CDS

ATT4D2 92611 bp DNA linear PLN 16-APR-2005
Arabidopsis thaliana DNA chromosome 3, BAC clone T4D2.

AL132958 GI:6434245

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 Nyakatura,G., Fartmann,B., Dauner,D., Sterr,W., Holland,R.,
Weichselgartner,M., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F.
and Salanoubat,M.
Unpublished

2 (bases 1 to 92611)

EU Arabidopsis sequencing,project.

Direct Submission

Submitted (21-DEC-1999) MIPS, at the Max-Planck-Institut fuer

Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Grouperment
d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr

Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.

QY 661 CGGTGTTCCATCGCTTCCAGAGGATATTTGCTGCTGATTTGGAGAGATATTCGCA 720
Db 512 CGGTGTTCCATCGCTTCCAGAGGATATTTGCTGCTGATTTGGAGAGATATTCGCA 571
QY 721 AATGTCAGAGAGGCTTTGTTGAGAGCTTGAGAGCTTTATGCAAGACGGAAGTGTGTCGA 780
Db 572 AATGTCAGAGAGGCTTTGTTGAGAGCTTGAGAGCTTTATGCAAGACGGAAGTGTGTCGA 631
QY 781 ATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGATCATTTGGTGTCTCTTTTAACT 840
Db 632 ATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGATCATTTGGTGTCTCTTTTAACT 691
QY 841 ACAATTTCCATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 692 ACAATTTCCATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 751
QY 901 AAGAAGCTGAGCTTCT 960
Db 752 AAGAAGCTGAGCTTCT 793
QY 961 TTTTCTGCTGTTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1020
Db 794 ----- 793
QY 1021 TTGTTTTTTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 794 -----AGATCGATGCTATTTGTTGATGATGATGATGATGATGATGATGATGATGAT 842
QY 1081 GCGTGTAACTTTGATAAGGGAACCTGTTGCAAGACCTTATGCTGATGTTGTTGTTGTTGTT 1140
Db 843 GCGTGTAACTTTGATAAGGGAACCTGTTGCAAGACCTTATGCTGATGTTGTTGTTGTTGTT 902
QY 1141 CGAAAGAGAGCGGAAGATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 1200
Db 903 CGAAAGAGAGCGGAAGATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 962
QY 1201 CTTGCGTTTCTGCTTGAAGCAGAGAGAGGTTTAACTTTTCTTAACTTTTCTTAACTTTTAA 1260
Db 963 CTTGCGTTTCTGCTTGAAGCAGAGAGG----- 991
QY 1261 CAAATTTGSCCTTTCATTATCGCAATAATCGAAGCTGATGTTGCTGAGGCTTTTTCAG 1320
Db 992 -----AG 993
QY 1321 GTATCAAGTGTGAGTACGAGTATCTTCTCAATGCTGTAGCGGGTCAATGAAC 1380
Db 994 GTATCAAGTGTGAGTACGAGTATCTTCTCAATGCTGTAGCGGGTCAATGAAC 1053
QY 1381 ACTGATCTGCTCTCACTTGGACACTTTATCTCTCAATGAGTACCTTAAATGATCTTC 1440
Db 1054 ACTGATCTGCTCTCACTTGGACACTTTATCTCTCAATG----- 1093
QY 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTACCTCTT 1500
Db 1094 ----- 1093
QY 1501 CAAATTTCTATGTTTATAGAACTCATCTGCAATTAAGGAAAGCAAGAGAGTAGACA 1560
Db 1094 -----AGAACTCATCTGCAATTAAGGAAAGCAAGAGAGTAGACA 1135
QY 1561 GAGTGTATAGAGGAAGAAACCGGCTTTCAGAGATATAAGAGGTTGAATGATCATCTC 1620
Db 1136 GAGTGTATAGAGGAAGAAACCGGCTTTCAGAGATATAAGAGGTTGAATGATCATCTC 1195
QY 1621 GTTGTATAAACGAGTCAATGCTCTATCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
Db 1196 GTTGTATAAACGAGTCAATGCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1236
QY 1681 TCTCTAATTAATTCATGAACATAAAATTTTCTGATGATGATGATGATGATGATGATGATGAT 1740
Db 1237 -----GTCTTGATTA 1246
QY 1741 GAAGAGCTCAAGTTCCTGACATTTCTTCTGTTGGAACTATAAGGTCAATACCGGACAAGACA 1800

Db 1247 GAAGAGCTCAAGTTCCTGACATTTCTTCTGGAACATATAAGGTCAATACCGGACAAGACA 1306
QY 1801 TTATGATTTTCAGTCTATACATCCATCGTTCTTCCGAGGTA 1841
Db 1307 TTATGATTTTCAGTCTATACATCCATCGTTCTTCCGAGGTA 1347

AY091083 1858 bp mRNA linear PLN 18-SEP-2002
Arabidopsis thaliana putative cytochrome P450 (At3g53130) mRNA,
partial cdb.
ACCESSION AY091083
VERSION AY091083.1 GI:20258841
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1858)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayaishizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1858)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayaishizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayaishizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J.,
Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGSC) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC)
contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source Location/Qualifiers
1. .1858
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAPL09-78-C14 (R19513)"
/ecotype="Columbia"


```
QY 781 ATATGGAAGCGAAGTCTCTCAGATGACACATTTGATGTCATTTGGGTTGCTCTCTTTTAACT 840
Db 636 ATATGGAAGCGAAGTCTCTCAGATGACACATTTGATGTCATTTGGGTTGCTCTCTTTTAACT 695
QY 841 ACAATTTCCGATCTTTTGACTACTGATAGTCTCTGTCATTTGAAGAGCTGTTTACACTGCTCTTA 900
Db 696 ACAATTTCCGATCTTTTGACTACTGATAGTCTCTGTCATTTGAAGAGCTGTTTACACTGCTCTTA 755
QY 901 AAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGAAGCAAGTTTCCCTGCTTT 960
Db 756 AAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGA----- 797
QY 961 TTTTCTGCTGTTGTTGATTTGTTGGCAACAATTTGAATCTTGTAAATTGAGAGGTTGG 1020
Db 798 ----- 797
QY 1021 TTGTTTTTTTTCAGATCGATGCAATTTGTAAAGATAGTCCCGACAGAGTGAAAGCTGAAAA 1080
Db 798 -----AGATCGATGCAATTTGTAAAGATAGTCCCGACAGAGTGAAAGCTGAAAA 846
QY 1081 GCCTGTAACTTTGATAAGGAAAACTGTTGAAGACCTTATTTGCTAAGTGTAAAGAAATGTT 1140
Db 847 GCCTGTAACTTTGATAAGGAAAACTGTTGAAGACCTTATTTGCTAAGTGTAAAGAAATGTT 906
QY 1141 CGAAAGAGAGCGCGAAGAAATCAATGATGAGGAGTATCTTAATGATGCTGACCCCAAGTAT 1200
Db 907 CGAAAGAGAGCGCGAAGAAATCAATGATGAGGAGTATCTTAATGATGCTGACCCCAAGTAT 966
QY 1201 CTTGCGTTTCTTCTGCTTGAAGCAGAGAGAGGTTTAAACTTTTTCCTTAAGTTTATAAG 1260
Db 967 CTTGCGTTTCTTCTGCTTGAAGCAGAGAG----- 995
QY 1261 CAAATTTTGGCCCTTTTCATTATCCGATPAATCGAAGCTGATTTGCTGATTTGAGGGTTTTCAG 1320
Db 996 -----AG 997
QY 1321 GPATCAAGTGTGCAAGTACCGGATGATCTTCTCTCAATGCTCTGAGCGGTCATGAACC 1380
Db 998 GPATCAAGTGTGCAAGTACCGGATGATCTTCTCTCAATGCTCTGAGCGGTCATGAACC 1057
QY 1381 ACTGATCTGCTCTCACTTGGACACTTATCTCTCAAGTAAAGTACCTTAAATGATATCTTC 1440
Db 1058 ACTGATCTGCTCTCACTTGGACACTTATCTCTCAAGTAA----- 1097
QY 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTACCTCTT 1500
Db 1098 ----- 1097
QY 1501 CAAATTTCTATGTTTCATAGAACTCATCTGCATTAAGGAAGCAACGAAGAGTAGACA 1560
Db 1098 -----AGAACTCATCTGCATTAAGGAAGCAACGAAGAGTAGACA 1139
QY 1561 GAGTGTAGAGGAAGAAACCGGCTTTTCAGGATATAAGAGTGAAGTACATCACTC 1620
Db 1140 GAGTGTAGAGGAAGAAACCGGCTTTTCAGGATATAAGAGTGAAGTACATCACTC 1199
QY 1621 GTTGTATAAACGAGTCAATGCGCTCTCATCTCATCTCTCTGTAAGCAATCAAGCTCATC 1680
Db 1200 GTTGTATAAACGAGTCAATGCGCTCTCATCTCATCTCTCT----- 1240
QY 1681 TCTTAATTTATCATGAACATAATTTTCTGATTTGATTTGTTCTTCTGTTAGTCTTTGATAA 1740
Db 1241 -----GTCCTTGATAA 1250
QY 1741 GAAGAGCTCAAGTTCTGACATTTCTTCTGGGAACTATAAGGTCAATACCGGACAGACA 1800
Db 1251 GAAGAGCTCAAGTTCTGACATTTCTTCTGGGAACTATAAGGTCAATACCGGACAGACA 1310
QY 1801 TTATGATTTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1841
Db 1311 TTATGATTTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1351
```

```
RESULT 5
ATHS26463 378 bp DNA linear PLN 29-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 120B11.
ACCESSION AJ526463
VERSION AJ526463.1 GI:26794723
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepiniec, B., Caboche, M. and Lecharny, A.
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
EMBO Rep. 3 (12), 1152-1157 (2002)
12446565
2 (bases 1 to 378)
Balzergue, S.
Direct Submision
Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
source
1..378
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="120B11"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Wassilewskija"
misc_feature
1..378
/note="T-DNA flanking sequence
left border"
ORIGIN
Query Match 15.3%; Score 376.4; DB 15; Length 378;
Best Local Similarity 99.7%; Pred. No. 2.4e-79; Indels 0; Gaps 0;
Matches 377; Conservative 0; Mismatches 1;
QY 858 ACTACTGATAGTCCCTGTCATTGAAGCTGTTTACACTGCTCTTAAAGAAAGCTGAGCTTCGT 917
Db 1 ACTACTGATAGTCCCTGTCATTGAAGCTGTTTACACTGCTCTTAAAGAAAGCTGAGCTTCGT 60
QY 918 TCTACTGATCTTCTGCGCATATTTGGAAGCGCAAGTTTCTGTTGTTTCTGTTGTTGTTG 977
Db 61 TCTACTGATCTTCTGCGCATATTTGGAAGCGCAAGTTTCTGTTGTTTCTGTTGTTGTTG 120
QY 978 ATTGTTGCGAACAAATTTGGATCTTGTAAATTCAGAGGGTTTGGTCTGTTTTTTCAGATCG 1037
Db 121 ATTGTTGCGAACAAATTTGGATCTTGTAAATTCAGAGGGTTTGGTCTGTTTTTTCAGATCG 180
QY 1038 ATGCAATTTGTGAAGATAGTCCCGACAGAGTCAAGCTGAAAGCGCTGAAAGCGCTGTAACTTGATAA 1097
Db 181 ATGCAATTTGTGAAGATAGTCCCGACAGAGTCAAGCTGAAAGCGCTGTAACTTGATAA 240
QY 1098 GGGAAACTCTTTGAAGACCTTATTTGCTAAAGAAATTTGTCGAAGAGAGGCGGAAA 1157
Db 241 GGGAAACTCTTTGAAGACCTTATTTGCTAAAGAAATTTGTCGAAGAGAGGCGGAAA 300
```


QY 1158 GAATCAATGATGAGGAGTATGAATGATGCTGACCCAAAGTATCTGCGCTTTCTGCTTG 1217
 |||||
 Db 301 GAATCAATGATGAAGAGTATGAATGATGCTGACCCAAAGTATCTGCGCTTTCTGCTTG 360
 |||||
 QY 1218 CAAGCAGACAGAGGTTTT 1235
 |||||
 Db 361 CAAGCAGACAGAGGTTTT 378
 |||||

RESULT 6
 AK220829
 LOCUS Arabidopsis thaliana gene for Cytochrom P450 -like protein, partial
 DEFINITION cds, clone: RAFL22-34-N18.
 ACCESSION AK220829
 VERSION AK220829.1 GI:62319016
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1
 AUTHORS Totoki Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
 Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
 Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
 Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1077)

AUTHORS Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
 Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
 Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
 Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
 Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE Direct Submission
 JOURNAL Submitted (22-MAR-2005) Motoaki Seki, RIKEN Genomic Sciences
 Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
 230-0045, Japan (E-mail: mseki@gsr.riken.jp).
 URL: http://range.gsc.riken.jp/, Tel: 81-45-503-9625,
 Fax: 81-45-503-9586

COMMENT An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al. (1998) Plant J. 15:707-720;
 Seki et al. (2002) Science 296:141-145).
 This clone is in a modified pBluescript vector.
 Please visit our web site (http://range.gsc.riken.jp/) for further
 details.

FEATURES
 source
 1..1077
 /location="Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL22-34-N18"
 /ecotype="Columbia"
 /note="common name: thale cress"

gene
 1..1077
 /gene="At3g53130"
 CDS
 1..906
 /gene="At3g53130"
 /codon_start=1
 /product="Cytochrom P450 -like protein"
 /protein_id="BAD94136.1"
 /db_xref="GI:62319017"
 /translation="NSPIEAVYITALKAEALRSITLLPYWKIDALCKIVPRQVRAEKA
 VTLREIVEDIYAKKEIVREGERINDEYVNDADPSILRFLASREVSQVLRDD
 LLSMLVAGHETGSLVTLVLLKNSALKAQEDVRVLEGRNPAPEDIKELKIIT
 RCINSEHLYPHEPVLIRAOVQPDILPGNYKVTQDQIMISVYNIHRSSEVWEKAEF
 LPERPDICGLPNEITDFKIFPSPGGPRKCGDQFALMEIALVALFLQRLNELVP
 DOTISMTTGATIHITNGLYMKVQSR"

ORIGIN
 Lycopersicon esculentum (Solanum lycopersicum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

Query Match 10.9%; Score 268.6; DB 15; Length 1077;
 Best Local Similarity 68.2%; Pred. No. 1.8e-53;
 Matches 557; Conservative 0; Mismatches 4; Indels 256; Gaps 3;
 QY 1025 TTTTTCAGATCGATGTCATTTGTAAAGATAGTCCCGAGACAGGTGAAAGCTCAAAAGGCT 1084
 |||||
 Db 73 TATGGAAGATCGATGTCATTTGTAAAGATAGTCCCGAGACAGGTGAAAGCTCAAAAGGCT 132
 |||||
 QY 1085 GTAACTTTTGAAGGAAACTGTTGAAGACCTTATTGTAAAGTAAAGAAATTTGCGAA 1144
 |||||
 Db 133 GTAACCTTTGATAAGGGAACCTGTTGAAGACCTTATTGTAAAGTAAAGAAATTTGCGAA 192
 |||||
 QY 1145 AGAAGGCGGAAAGAAATCAATGATGAGGATATGTAATGATGCTGACCCAAAGTATCTCTG 1204
 |||||
 Db 193 AGAAGGCGGAAAGAAATCAATGATGAGGATATGTAATGATGCTGACCCAAAGTATCTCTG 252
 |||||
 QY 1205 CGTTTCTTGTCTTGAAGCAGAGAGGTTTAAACTTTTTCCTTAAAGTTTATAAGCAA 1264
 |||||
 Db 253 CGTTTCTTGTCTTGAAGCAGAGAGGTTTAAACTTTTTCCTTAAAGTTTATAAGCAA 277
 |||||
 QY 1265 TTTGGCCTTTTCATTTATCGCATAATCGAAGCTGATGTTGCATTTGTGAGGGTTTTTCAGG 1324
 |||||
 Db 278 -----AGGTAT 283
 |||||
 QY 1325 CAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCAATGAACACCTG 1384
 |||||
 Db 284 CAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCAATGAACACCTG 343
 |||||
 QY 1385 GATCTGCTCTCACTGGACACTTTATCTCTAAAGTAAAGTACCTTAAATGATCTTCTACT 1444
 |||||
 Db 344 GATCTGCTCTCACTGGACACTTTATCTCTCTTAAAGTTA----- 379
 |||||
 QY 1445 TTGCTATGCTAGAGAAATTTTACTTGGATGGAGCTTCTCTGTTCTCATTTTACCTCTTCAA 1504
 |||||
 Db 380 ----- 379
 |||||
 QY 1505 TTCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGACACAAGAAAGTAGACAGAT 1564
 |||||
 Db 380 -----AGAACTCATCTGCAATTAAGGAAGACACAAGAAAGTAGACAGAT 425
 |||||
 QY 1565 GTTAGAAGGAAGAACCGCGCTTTTCGAGGATATAAAGAGTTGAAGTACATCACTCGTTG 1624
 |||||
 Db 426 GTTAGAAGGAAGAACCGCGCTTTTCGAGGATATAAAGAGTTGAAGTACATCACTCGTTG 485
 |||||
 QY 1625 TATAAACGAGTCAATGCGCTCTATCTCTCATCTCTGTAAGCAATCAAGCTCATCTCTC 1684
 |||||
 Db 486 TATAAACGAGTCAATGCGCTCTATCTCTCATCTCTCT----- 522
 |||||
 QY 1685 TAATTTATTCATGAATAAAATTTTCGATTGATTTGTTCTCTGTTAGGTTTGATGAAGAAG 1744
 |||||
 Db 523 -----GTCTTGATAAGAAG 536
 |||||
 QY 1745 AGCTCAAGTTCCTGACATTTCTCTGGGAACCTATAGGTCAATACCGGACAGACATTAT 1804
 |||||
 Db 537 AGCTCAAGTTCCTGACATTTCTCTGGGAACCTATAGGTCAATACCGGACAGACATTAT 596
 |||||
 QY 1805 GATTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1841
 |||||
 Db 597 GATTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 633
 |||||

RESULT 7
 BT012891
 LOCUS Lycopersicon esculentum clone 113997F, mRNA sequence.
 DEFINITION BT012891
 ACCESSION BT012891.1 GI:47104306
 VERSION FLI CDNA.
 KEYWORDS Lycopersicon esculentum (Solanum lycopersicum)
 SOURCE Lycopersicon esculentum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

AP008216_036 3600001 3710000
AP008216_037 3700001 3810000
AP008216_038 3800001 3910000
AP008216_039 3900001 4010000
AP008216_040 4000001 4110000
AP008216_041 4100001 4210000
AP008216_042 4200001 4310000
AP008216_043 4300001 4410000
AP008216_044 4400001 4510000
AP008216_045 4500001 4610000
AP008216_046 4600001 4710000
AP008216_047 4700001 4810000
AP008216_048 4800001 4910000
AP008216_049 4900001 5010000
AP008216_050 5000001 5110000
AP008216_051 5100001 5210000
AP008216_052 5200001 5310000
AP008216_053 5300001 5410000
AP008216_054 5400001 5510000
AP008216_055 5500001 5610000
AP008216_056 5600001 5710000
AP008216_057 5700001 5810000
AP008216_058 5800001 5910000
AP008216_059 5900001 6010000
AP008216_060 6000001 6110000
AP008216_061 6100001 6210000
AP008216_062 6200001 6310000
AP008216_063 6300001 6410000
AP008216_064 6400001 6510000
AP008216_065 6500001 6610000
AP008216_066 6600001 6710000
AP008216_067 6700001 6810000
AP008216_068 6800001 6910000
AP008216_069 6900001 7010000
AP008216_070 7000001 7110000
AP008216_071 7100001 7210000
AP008216_072 7200001 7310000
AP008216_073 7300001 7410000
AP008216_074 7400001 7510000
AP008216_075 7500001 7610000
AP008216_076 7600001 7710000
AP008216_077 7700001 7810000
AP008216_078 7800001 7910000
AP008216_079 7900001 8010000
AP008216_080 8000001 8110000
AP008216_081 8100001 8210000
AP008216_082 8200001 8310000
AP008216_083 8300001 8410000
AP008216_084 8400001 8510000
AP008216_085 8500001 8610000
AP008216_086 8600001 8710000
AP008216_087 8700001 8810000
AP008216_088 8800001 8910000
AP008216_089 8900001 9010000
AP008216_090 9000001 9110000
AP008216_091 9100001 9210000
AP008216_092 9200001 9310000
AP008216_093 9300001 9410000
AP008216_094 9400001 9510000
AP008216_095 9500001 9610000
AP008216_096 9600001 9710000
AP008216_097 9700001 9810000
AP008216_098 9800001 9910000
AP008216_099 9900001 10010000
AP008216_100 10000001 10110000
AP008216_101 10100001 10210000
AP008216_102 10200001 10310000
AP008216_103 10300001 10410000
AP008216_104 10400001 10510000
AP008216_105 10500001 10610000
AP008216_106 10600001 10710000
AP008216_107 10700001 10810000
AP008216_108 10800001 10910000

AP008216_109 10900001 11010000
AP008216_110 11000001 11110000
AP008216_111 11100001 11210000
AP008216_112 11200001 11310000
AP008216_113 11300001 11410000
AP008216_114 11400001 11510000
AP008216_115 11500001 11610000
AP008216_116 11600001 11710000
AP008216_117 11700001 11810000
AP008216_118 11800001 11910000
AP008216_119 11900001 12010000
AP008216_120 12000001 12110000
AP008216_121 12100001 12210000
AP008216_122 12200001 12310000
AP008216_123 12300001 12410000
AP008216_124 12400001 12510000
AP008216_125 12500001 12610000
AP008216_126 12600001 12710000
AP008216_127 12700001 12810000
AP008216_128 12800001 12910000
AP008216_129 12900001 13010000
AP008216_130 13000001 13110000
AP008216_131 13100001 13210000
AP008216_132 13200001 13310000
AP008216_133 13300001 13410000
AP008216_134 13400001 13510000
AP008216_135 13500001 13610000
AP008216_136 13600001 13710000
AP008216_137 13700001 13810000
AP008216_138 13800001 13910000
AP008216_139 13900001 14010000
AP008216_140 14000001 14110000
AP008216_141 14100001 14210000
AP008216_142 14200001 14310000
AP008216_143 14300001 14410000
AP008216_144 14400001 14510000
AP008216_145 14500001 14610000
AP008216_146 14600001 14710000
AP008216_147 14700001 14810000
AP008216_148 14800001 14910000
AP008216_149 14900001 15010000
AP008216_150 15000001 15110000
AP008216_151 15100001 15210000
AP008216_152 15200001 15310000
AP008216_153 15300001 15410000
AP008216_154 15400001 15510000
AP008216_155 15500001 15610000
AP008216_156 15600001 15710000
AP008216_157 15700001 15810000
AP008216_158 15800001 15910000
AP008216_159 15900001 16010000
AP008216_160 16000001 16110000
AP008216_161 16100001 16210000
AP008216_162 16200001 16310000
AP008216_163 16300001 16410000
AP008216_164 16400001 16510000
AP008216_165 16500001 16610000
AP008216_166 16600001 16710000
AP008216_167 16700001 16810000
AP008216_168 16800001 16910000
AP008216_169 16900001 17010000
AP008216_170 17000001 17110000
AP008216_171 17100001 17210000
AP008216_172 17200001 17310000
AP008216_173 17300001 17410000
AP008216_174 17400001 17510000
AP008216_175 17500001 17610000
AP008216_176 17600001 17710000
AP008216_177 17700001 17810000
AP008216_178 17800001 17910000
AP008216_179 17900001 18010000
AP008216_180 18000001 18110000
AP008216_181 18100001 18210000

AP008216_182	18200001	18310000				
AP008216_183	18300001	18410000				
AP008216_184	18400001	18510000				
AP008216_185	18500001	18610000				
AP008216_186	18600001	18710000				
AP008216_187	18700001	18810000				
AP008216_188	18800001	18910000				
AP008216_189	18900001	19010000				
AP008216_190	19000001	19110000				
AP008216_191	19100001	19210000				
AP008216_192	19200001	19310000				
AP008216_193	19300001	19410000				
AP008216_194	19400001	19510000				
AP008216_195	19500001	19610000				
AP008216_196	19600001	19710000				
Query Match 9.7%; Score 239.6; DB 15; Length 110000;						
Best Local Similarity 63.2%; Pred. No. 9.8e-47;						
Matches 425; Conservative 0; Mismatches 204; Indels 43; Gaps 2;						
QY	605	ATCTCTCTATTGGGCTTAGCTGAAGTCTGATTTTTTACATAGGCGAGCGGTAGACGGT	664			
Db	50743	ATAATCTCTGTTTCAACTTTACACTTGTGTTTGTGATCTGTTTCAGGTGACACGTCGATCAGT	50802			
QY	665	GGTTCATCGCTTCACAGAGGATTTTGTCTGTGATTTGGAGAGATTTCTGCAATG	724			
Db	50803	TGTACCATCTCTACAAACGATTTCTCTCGGTGATGTTGACAGATTTTTTGTAAATG	50862			
QY	725	TGCAGAGAGCTTGTGTGAGAGTTGCAGGCTTTATGCAGAAAGCGGAAGTCTGTGAATAT	784			
Db	50863	TGCTGAGAGATTTAGTGGAGAGCTTTGAGACATCTGCTTTAAGTGCACAACTGTAATAT	50922			
QY	785	GGAAGCGAAGTTCTCTCAGATGACACTTGATGTCATTTGGTTGTCTCTTTTTTAACACAA	844			
Db	50923	GGAAGCGAAGTTCTCTCAAATGACTTTTAGATGATGTTGTTTGTCTTGTTCATTTACAA	50982			
QY	845	TTTTCGATCTTTTGACTACTGATGATGCTCTGTCATTCGAAGCTGTTTACACTGCTCTTAAAGA	904			
Db	50983	TTTTGTATCCCTCACATCAGATAGCCCTGTTATGATGCTGTTTACACTGCACCTCAGGA	51042			
QY	905	AGCTGAGCTTCGTTCTACTGATCTTCTGTCATATTGGAAGCGCAAGTTTCC-----TG	956			
Db	51043	AGCAGAACTTCGTTCTACAGATCTTTTACCATACTGGAAGGTAATGTTCTCTTTGCCATG	51102			
QY	957	TGTTTTCTCTGTTGTTGTTGATGTTGTTGGAACAAATGGAATCTCTGTTAAATGA-----	1010			
Db	51103	TTATTTTCTAAAATAATCCAAATATGTTGTCGCTATCTAACTTCTAGTGAAGCAAAATGC	51162			
QY	1011	-----GAGGGTTTGGTTGTTTCTGATGTAGCCCTTGTGTTCTTGTAGATGTTGATTT	1041			
Db	51163	TGACTGATGCTGCTGACTGATGTTGATGTAGCCCTTGTGTTCTTGTAGATGTTGATTT	51222			
QY	1042	ATTGTGTAAGATAGTCCCGACAGAGTGAAGCTGAAAAGGCTGTAACCTTTGATAGGGA	1101			
Db	51223	GCTGTGCAAGATTTGTTCTCTAGACAAATAAAGCAGAAAGGCGATTAACATCATCAGAA	51282			
QY	1102	AACTGTTGAAGACCTTATTGCTTAAGTGTAAAGAAATTTGCAAGAGAGGCGCAAGAT	1161			
Db	51283	TACCGTTGAGGACCTAATTTACCAATGCAAGAAGATTTGTAGATGCTGAGAATGAACAAAT	51342			
QY	1162	CAATGATGAGGATGATGTAATGATGCTGACCCCAAGTATCCTGCGTTTTCTGCTTGCAG	1221			
Db	51343	TGAGGGTGAGGAATATGTAATGAGGACACCCCTAGCATCTCTGCGATTCCTACTTCTAG	51402			
QY	1222	CAGAGAAGAGGT	1233			
Db	51403	CCGTGAAGAGGT	51414			

RESULT 9
AC025783
LOCUS
DEFINITION Oryza sativa chromosome 10 BAC OSJNBa0001014 genomic sequence,
AC025783 178022 bp DNA linear PLN 23-MAY-2002

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

complete sequence.
AC025783
AC025783.10 GI:21104901
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 178022)
Buell, C.R., Yuan, Q., Ouyang, S., Moffat, K.S., Hill, J.N., Burr, P.C.,
Hsiao, J., Ziemann, V., Pai, G., Bowman, C.D., Fujii, C.Y., Khalak, H.,
VanAken, S.E., Bowman, C.L., Craven, B., Utterback, T.R., Khalak, H.,
Feldblyum, T.V., Quackenbush, J., White, O., Salzberg, S.L. and
Fraser, C.M.
Oryza sativa chromosome 10 BAC OSJNBa0001014 genomic sequence
Unpublished
2 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (14-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (14-MAR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (16-MAR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
5 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (28-AUG-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
6 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (25-APR-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
7 (bases 1 to 178022)
Buell, R.
Direct Submission
Submitted (23-MAY-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
On May 23, 2002 this sequence version replaced gi:20303636.
Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0001014 is from Oryza sativa chromosome 10
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan and Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkEM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), Fgenesh
Steven Salzberg, contact mperte@tigr.org), searches of the
complete sequence against a peptide database and the plant EST
database at TIGR (http://www.tigr.org/tdb/tgi.shtml). Annotated
genes are named to indicate the level of evidence for their
annotation. Genes with similarity to other proteins are named after
the database hits. Genes without significant protein similarity but
with EST similarity are named as unknown proteins. Genes without
protein or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

1331 bp sequence identical to E. coli transposase GB:S67119

GI:455674 has been cut from this BAC.

This BAC overlaps with rice BAC OSJNBa0063E23 (AC113947) and OSJNBa0082M15 (AC020666).

FEATURES

source

Location/Qualifiers

1. .178022
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
/map="near C809"
/clone="OSJNBa0001014"
/note="Japonica cultivar-group"
660. .1022
/gene="OSJNBa0001014.22"
/note="predicted by genescan"
660. .1022
/gene="OSJNBa0001014.22"
660. .1022
/gene="OSJNBa0001014.22"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK20056.1"
/db_xref="GI:13357259"
/translation="MFPNPMGCGPFRSTRQPLRHGGGGVTVVSRGGGGGAA
RKIVVGSDELRLIAAGVTRQCAVSGSPPELLQQQARATEVAAPMPRRPSEGGALV
AGRRGEQFALDGIPEEA"
766. .786
/rpt_family=" (CGG)n"
3343. .3469
/rpt_family="putative Stowaway type MITE element"
complement(3354. .3480)
/rpt_family="putative Stowaway type MITE element"
complement(3808. .3839)
/rpt_family="AT rich"
complement(3937. .3963)
/rpt_family="AT rich"
complement(4838. .4871)
/rpt_family=" (TA)n"
4845. .4870
/note="Overlap BAC OSJNBa0063E23 (AC113947) has 17 AT
repeats in this region."
5433. .5567
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
5441. .10390
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
5499. .5636
/rpt_family="putative Tourist type MITE element"
complement(5712. .6743)
/rpt_family="AT rich"
complement(9156. .9226)
/rpt_family="retroelement from gi|5295936|AB026295 Oryza
sativa (160208 to 165872) 5665 nt"
10278. .10452
/rpt_family="Tourist-like MITE element from O.sativa DNA
for LRK2 gene gi|1263159|emb|X89226.1|OSLRK2GEN (1 to 540)
540 nt"
complement(10291. .10434)
/rpt_family="Tourist OLI MITE element from gb:U72728 Oryza
longistaminata receptor-like kinase protein (Xa21), family
member F, pseudogene sequence (1080 to 1236) 157 nt"
10300. .10397
/rpt_family="putative Tourist type MITE element"
10324. .10474
/rpt_family="putative Tourist type MITE element"
complement(12414. .12457)
/rpt_family="AT rich"
15036. .17810
gene

mrna
/gene="OSJNBa0001014.25"
/note="predicted by genescan"
join(<15036. .15183,15255. .15433,16442. .16525,17234. .17401,
17688. .>17810)
/gene="OSJNBa0001014.25"
join(15036. .15183,15255. .15433,16442. .16525,17234. .17401,
17688. .17810)
/gene="OSJNBa0001014.25"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK20044.1"
/db_xref="GI:13357247"
/translation="MDASPSSSSSPVAPEVQWAGGAGGSGAIGPACRVLCVAT
CVFAVGVGAVTSGAIGLATESKMGKAGIGAGVFSIEVBSRDLMHSSDSA
VWCLVYMDYIISSLSSGLRVKPGVAVSQISAISSPFASTSLFETGGTGL
PADTLRLPAIKITGDNAVDVSGEPCICSCVCLQDFRGMARLPSCRHVFHVPICDC
WLVRHGSCPLCRRII"
complement(20477. .20596)
/rpt_family="Gaijin_OEI MITE element from gb:X52623 Rice
4-CL gene for 4-coumarate-CoA ligase (1888 to 2027) 140
nt"
22465. .22516
/rpt_family=" (GAAA)n"
complement(23182. .23375)
/rpt_family="Rice RS0881X19 insert DNA with repeat units
RI-R6 gi|20362|emb|X55420.1|OSRSRIK6"
complement(23216. .23353)
/rpt_family="DNA fragment containing the junction of AA
genome-specific tandem repeat sequence TrsA
(telomere-associated) gi|804936|dbj|D16452.1|RICTRSA1 Rice
(O.sativa), TrsA1 locus"
complement(23490. .23515)
/rpt_family="AT rich"
complement(23739. .23847)
/rpt_family=" (CGG)n"
23861. .24031
/rpt_family=" (CCGCG)n"
complement(23942. .24112)
/rpt_family=" (CGG)n"
complement(24108. .24214)
/rpt_family=" (AGGGG)n"
complement(24234. .24404)
/rpt_family=" (CGG)n"
24521. .24618
/rpt_family=" (GAAAA)n"
complement(24651. .24680)
/rpt_family="AT rich"
complement(27603. .27652)
/rpt_family="GC rich"
27711. .27833
/rpt_family=" (CCGCG)n"
complement(28641. .29130)
Query Match 9.7%; Score 239.6; DB 15; Length 178022;
Best Local Similarity 63.2%; Pred.No. 9.3e-47;
Matches 425; Conservative 0; Mismatches 204; Indels 43; Gaps 2;
QY 605 ATCTTCTATTGGGCTTAGCTGAAGTCTGATTTTTTACATAGCGGCGGTAGAGCGGT 664
DB 113142 ATATCTCTTCTCAACTTACACTTGTGTTTGGTATCTGTTCAAGTGAGACGTGATCAGT 113201
QY 665 GGTTCATCGCTTCAAGAGGATTTCTCTGTGATTTGTGAGAGAGATTTCTGCAATG 724
DB 113202 TGTACCATCTCTACACAAACGATTTCTCTCGTGATGTTGACAGAGTTTTTTGTAATG 113261
QY 725 TGCAGAGGCGTCTTGTGAGAGTTCCGCCCTTATGCAGAACGACGAGTGTGCAATAT 784
DB 113262 TGCTGAGAGATTAGTGGAGAAAGCTTGAGACATCTGCTTTAAGTGCAACCTGTAAATAT 113321
QY 785 GGAAGCGAAGTTCTCTCAGATGACACTTGATGTCACTGGGTGTCTCTTTTAACTACAA 844
DB 113322 GGAAGCAGGTTCTCTCAATGACTTTAGATGTGATGGTTGTTCTTCTTCAATTACAA 113381

Qy	845	TTTCGATCTTTGACTACTGATGATCCCGTCATTTGAAGCTGTTTACACTGCTCTTAAAGA	904
Db	113382	TTTTGATTCCTTCACATCAGATAGCCCTGTTATTGATGCTGTTTACACTGCACTCAAGGA	113441
Qy	905	AGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTGGAAGCGCAAGTTTCC-----TG	956
Db	113442	AGCAGACTTCTTCTACAGATCTTTTACCATACTGGAAGGTAATGTTCTTTTGGCCATG	113501
Qy	957	TGTTTTTCTGTGGTTGTTGATGTGTGGAACTATGGAATGATCTTGTGTAATGA-----	1010
Db	113502	TTATTTTCTAAAAAATCCAAATTATGTTGGCTATCTAACTTCTAGTCAGACGAAATTGC	113561
Qy	1011	-----GAGGGTTTGGTTGTTTCTTCTAGATGATGATTT	113621
Db	113562	TGACTGATGTCGCTGACTGATGTCGTGTAGTGTAGCCCTTGTGTTCTTCTGATGATGATTT	113681
Qy	1042	ATTGTTGAAGATAGTCCCGAGACAGCTGAAAGCTGAAAAGGCTGTGAACTTTGATAAGGGA	1101
Db	113622	GCTGTGCAAGATTGTCTTAGCAATAAAGCAGAAAAGCGAGTTACATCATCAGGAA	113681
Qy	1102	AACTGTTGAAGACCTTATTGCTTAAGTGAAGAAATGTCGAAAGAGAGGCGAAAGAAAT	1161
Db	113682	TACCGTTGAGGACCTAAATTACCAATGCAAGAAGATTGTAGATGCTGAGAATGAACAAAT	113741
Qy	1162	CAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTATCCTCGCTTCTTCTGCTGCAAG	1221
Db	113742	TCAGGCTGAGGAATATGTAATGAGGCAGACCCCTAGCATCTCTGCGATTCTTACTTGCTAG	113801
Qy	1222	CAGACAAAGAGGT	1233
Db	113802	CCGTGAAGAGGT	113813
RESULT 10			
LOCUS	AE017117	299735 bp	DNA linear PLN 06-JUN-2003
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 10, section 71 of 77 of the complete sequence.		
ACCESSION	AE017117	AE016959	
VERSION	AE017117.1	GI:31433358	
KEYWORDS	Oryza sativa (japonica cultivar-group)		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 299735)		
AUTHORS	The Rice Chromosome 10 Sequencing Consortium		
CONSRTM	In-depth view of structure, activity, and evolution of rice chromosome 10		
TITLE	Science 300, 1566-1569 (2003)		
JOURNAL	2 (bases 1 to 299735)		
REFERENCE	Buell, C.R., Wing, R.A., McCombie, W.R., Messing, J. and Yuan, Q.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (05-MAY-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).		
FEATURES	Location/Qualifiers		
	1..299735		

misc_feature	
gene	
mRNA	
CDS	
gene	
mRNA	
CDS	
gene	
mRNA	
CDS	

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultiivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
<1..133260
/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSJNBa0001014 (GB:AC025783)."
<431..>3685
/locus_tag="OSJNBa0001014.2"
/note="EST D48843, BE230623 from this gene"
join(431..633,699..781,864..905,1063..1104,1442..1525,1702..1789,1913..2010,3644..>3685)
/locus_tag="OSJNBa0001014.2"
join(471..633,699..781,864..905,1063..1104,1442..1525,1702..1789,1913..2010,3644..3685)
/locus_tag="OSJNBa0001014.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP54878.1"
/db_xref="GI:31433359"
/translation="MAMAAATPVLLLRSTILLPRSPRASACVLAPORRGRRVRLRSPPPELPGELMEEDSKFVPLNAEDPMYPPALLLIGFEETGNKVQFLRLDGEFLKVIHCTEEMTKQLWDAMHTEQPSIEAVKIAKSMPRCIFSGLTGEEMMFINSYPTEGLEPAAFAALVPNSADKILGEVIEEINGDHEMLLLVAGDFSRSCYC"<6954..>9818
/locus_tag="OSJNBa0001014.3"
/note="Similar to C 4 sterol methyl oxidase GB:AAC49139 GI:1161339 (Saccharomyces cerevisiae); EST D15076, C97849 from this gene"
join(6954..7365,7735..8057,9074..9221,9316..9348,9516..9818)
/locus_tag="OSJNBa0001014.3"
join(7036..7365,7735..8057,9074..9221,9316..9348,9516..9587)
/locus_tag="OSJNBa0001014.3"
/codon_start=1
/product="putative C-4 sterol methyl oxidase"
/protein_id="AAP54879.1"
/db_xref="GI:31433360"
/translation="MLPYATAEAALGRAMTAABSLWFRYSAGIDPYLVFWNHLF LFWFTLAPLVALLEAPAAVGPFLQPKVLRSEFFCYCDVRMLFLVLIGPLQ LVSPYTVQVGHITGLPLPSLGEAAQLVFLVEDYLNNTYHRLHGSWGYEKIHRV HFEFTAPGFAAPYAHVAEVLILGIPSPVGPALAPGHMTTWLWTLVLRQMEAIEHSG FDPFNLTKYIPFYCGAAYHYHYHVGVSQSNFASVFTCYDLYGTDKGYYHKAYQ AKMKALQTEGEKADSNGLSYAKLD complement(<10045..>13873)
/locus_tag="OSJNBa0001014.4"
/note="contains pfam profile PF00098 (zinc finger, CCHC class); EST AU161993, AU057223, AU093006, AU057224 from this gene"
complement(join(<10045..10997,11698..11793,12830..13057,13670..13873))
/locus_tag="OSJNBa0001014.4"
complement(join(10454..10997,11698..11793,12830..12933))
/locus_tag="OSJNBa0001014.4"
/codon_start=1
/product="putative zinc finger protein"
/protein_id="AAP54880.1"
/db_xref="GI:31433361"
/translation="MSSRSPPKRIIRRTERTSYRDAPYRDSRRGSPFPNDLCNCK KEPGHFARDCPNVALCHACGLPGHIAECSSKOLCWNCKEPGHMANSCNEGICRNG KSGHIARECSAPPMLPGEMRLCSNCKYKPGHLAECTNEKACNNCRKSGHLARNCPNEP VCNLNVSGHLARECPKSDAINEGRGYPFRGYSDVVCRACNQVGHMSRDCWAGAFM ICHNCGRGHMAYECPGSLMDRFPFRRR"<14813..>17815
/locus_tag="OSJNBa0001014.12"
/note="predicted by fgenesh"
join(<14813..15352,16307..16418,16786..16950,17391..>17815)
/locus_tag="OSJNBa0001014.12"
join(14813..15352,16307..16418,16786..16950,17391..17815)

DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	Oryza sativa (japonica cultivar-group) cDNA clone:J013039G17, full insert sequence.	
	AK065689 1 GI:32975707	
	FLI_CDNA; CAP trapper.	
	Oryza sativa (japonica cultivar-group)	
	Oryza sativa (japonica cultivar-group)	
REFERENCE AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	
	1	
	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team., Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuka,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN: Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayaehizaki,Y.	
	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice	
	Science 301 (5631), 376-379 (2003)	
TITLE JOURNAL PUBMED REFERENCE AUTHORS	12869764	
	2 (bases 1 to 1876)	
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Hayashi,S., Hanagaki,T., Hara,A., Hashidume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagiwara,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koyama,S., Kurihara,C., Kuroski,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.	
	Direct Submission	
	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)	
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice.	
	URL : http://cdna01.dna.affrc.go.jp/cDNA/	
	NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.	
	PAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kouda,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.	
	10	

FEATURES source	Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagiwara,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Oota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayaehizaki,Y.	
	Location/Qualifiers	
	1..1876	
	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="mRNA"	
ORIGIN	/culivar="Nipponbare"	
	/db_xref="taxon:39947"	
	/clone="J013039G17"	
	Query Match 7.4%; Score 183; DB 15; Length 1876;	
	Best Local Similarity 61.8%; Pred. No. 6.6e-33;	
QY	Matches 363; Conservative 0; Mismatches 135; Indels 89; Gaps 1;	
	647 GCGCAGGCGTAGAGCGGTGGTCCATCGCTTACAGGAGGTATTGTCTGTGATGTGGA 706	
	626 GGTGAGAGCTCGATCAGTTGTACCATCTCTACACAAACGATTTCTCTCGGTGATGTGA 685	
	707 GAGAGTATTCGCAAAATGTGCAGAGAGCGCTTGTGTGGAAGTTGCGAGCCTTATGCAGAGA 766	
	686 CAGAGTTTTTTTGTAAATGTCTGAGAGATTAGTGGAGAAGCTTGAGACAATCTGCTTTAAG 745	
QY	767 CGGAGTGTCTGTGAATATGGAAGCGAAGTCTCTCAGATGACACATTTGATGCTGGGTT 826	
	746 TGGCAAACTGTAAATATGGAAGCAAGGTCTCTCAATGACTTTAGATGTGATTTGGTTT 805	
	827 GTCTCTTTTAACTAACAATTCGATCTTTTGACTACTGATAGTCTCTGTCAATTTGAAGCT 886	
	806 GTCTCTTTTCAATTACAATTTTGAATTCCTCACAATCAGATAGCCCTGTTATTGATGCTGT 865	
	887 TTACACTGCTCTTAAAGAGCTGAGCTTGTCTACTACTGATCTTCTGCCATATTTGGAAGGC 946	
QY	866 TTACACTGCACTCAAGGAAGCAAGAACTTCGTCTCAAGATCTTTTACCATCTGGA---- 921	
	947 AAGTTTCTCTGTGTTTTTTTCTGTGGTTTTTGTGATTGTGTGGAAACAATTTGGATTCTTGT 1006	
	922 ----- 921	
	1007 TTGAGAGGGTTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGTGAGTAGTCCCGAGACAG 1066	
	922 -----AGATTGATTTGCTGTGCAAGATTGTTCTCTAGACAA 956	
QY	1067 GTGAAAGCTGAAAAGCGTGTAACTTTGTAAGGAAGAACTGTTGAAGACCTTATTGTCTAAG 1126	
	957 ATAAAAGCAGAAAAGGCGAGTTAAACATCATCAGGAATACCGTTGAGGAGCTTAATTACCAA 1016	
	1127 TGTAAAGAAATTTGTGCAAAAGAGAGGCGCAAGAAATCAATGATGAGGAGGTATGTAATGAT 1186	
	1017 TGCAGAGAAATTTGTAGATCTGAGATGAACAAATTTGAGGGTGAAGGAATATGTAATGAG 1076	
	1187 GCTGACCCCAAGTATCTCGGTTTCTTGTGTTTGAAGCAGAGAGAGGT 1233	
QY	1077 GCAGACCTTAGCATCTCGGATTCCTTACTTCTTGTAGCCGTGAAGAGGT 1123	
	RESULT 12	
	AC155564	
	LOCUS	
	Zea mays strain B73 clone ZMWBc0171N08, *** SEQUENCING IN PROGRESS	

AC155564 205408 bp DNA linear HTG 25-JAN-2005
DEFINITION Zea mays strain B73 clone ZMWBc0171N08, *** SEQUENCING IN PROGRESS

/estimated_length=unknown
82989. .83088
gap /estimated_length=unknown
84258. .84357
gap /estimated_length=unknown
88369. .88468
gap /estimated_length=unknown
103421. .103520
gap /estimated_length=unknown
111901. .112000
gap /estimated_length=unknown
139515. .139614
gap /estimated_length=unknown
141413. .141512
gap /estimated_length=unknown
146354. .146453
gap /estimated_length=unknown
150359. .150498
gap /estimated_length=unknown
152461. .152560
gap /estimated_length=unknown
153780. .153879
gap /estimated_length=unknown
167813. .167912
gap /estimated_length=unknown
170603. .170702
gap /estimated_length=unknown
198609. .198708
gap /estimated_length=unknown
201018. .201117
gap /estimated_length=unknown
204280. .204379
gap /estimated_length=unknown

ORIGIN

Query Match 6.0%; Score 146.8; DB 14; Length 205408;
Best Local Similarity 65.0%; Pred. No. 1.9e-24;
Matches 217; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 632 CTGATTATTTTACATAGCGAGCGGTAGACGGTGTTCATCGTTCACAGGAGGTATT 691
Db 108385 CTGGTATGTAATTAGTGTGAGACGCAAGTGTGTGCTTCATTGTCATAAAATATTTAT 108444

QY 692 GTCTGTGATGTGGAGAGATATCTGCAATGTGACAGAGCGTGTTCAGAGTTGCA 751
Db 108445 GTCCATTATCGTGAAGAGTATTTTGTAAATATGATGAGAGATTAATAGACAACTTCA 108504

QY 752 GCCTTATGACAGACGGAAGTGTGTAATATGGAAGCGAAGTTCTCTCAGATGACACT 811
Db 108505 GCCTATGCTTTGAGTGGGGAACCTGTCAATATGAAGCAAGTTTTCTCAGTTGACATT 108564

QY 812 TGATGTCAATGGTGTCTCTTTTAACTACAATTTTCGATTTCTTTGACTACTGATGTC 871
Db 108565 GGATGTATTTGTTTGTTCATTTGTTCACTACACTTTTATTCCTCAGATGACAGTTTC 108624

QY 872 TGTCAATGAGCTGTTTACATGCTCTTTAAAGAGCTGAGTTCGTTCTACTGATCTTCT 931
Db 108625 TAGCATTTGTTGTTTATCTACTGCACTCAAGAAGCAGAGATTCGTTCTACAGATCTTT 108684

QY 932 GCCTATTTGGAAGCAAGTTTCTCTGTTGTTTTTC 965
Db 108685 TGCCATACTAGAAGTGTGCTCTCTTTAACTTC 108718

RESULT 13

AY142017
LOCUS AY142017 1788 bp mRNA linear PLN 02-SEP-2002
DEFINITION Arabidopsis thaliana Atg31800/68069_m00159 mRNA, complete cds.
ACCESSION AY142017
VERSION AY142017.1 GI:22655377
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1788)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Heuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 1788)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Heuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

TITLE
JOURNAL

Submitted (15-AUG-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs (Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.J., Dale, J.M., Deng, J.M., Heuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1. .1788
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U14026"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

CDS

1. .1788
/note="putative cytochrome P450"
/codon_start=1
/product="Atg31800/68069_m00159"
/protein_id="AA098281.1"
/db_xref="GI:22655378"

ORIGIN

Query Match 3.7%; Score 91; DB 15; Length 1788;
 Best Local Similarity 60.9%; Pred. No. 9e-11;
 Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGATTCTGCAATGTGCAGAGAGGCTTGTGAGAGTTGCAGCCTTATGC 760
 DB 645 TATGATTATTTATTCGGAGAGCTTCAGATAGGCTTTGTTCAGAGCTTGATGCTGCTC 704

QY 761 AGAAGACGGAAGTGTGTAATATGAAGCGAAGTCTCTCAGATGACACTTGTATCAT 820
 DB 705 ATTGAAGGCGAAGATAGAGATGAATCACTCTCTCTCTGTTGACACTTGATATTAT 764

QY 821 TGGGTGTCTCTTTTAACTACAATTCGATCTTTTGACTACTGATGCTGCTCATGCA 880
 DB 765 TGGCAAGGCGGTTTTCAATTACGACTTTCGACTCCCTTACTATGATACCGGTTGATCGA 824

QY 881 AGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCATATG 940
 DB 825 GCGAGTACACTGTTCTTAAGAGAGCTGAGACAGAGTGTTCACCTATTCCTGTTG 884

QY 941 GAA 943
 DB 885 GCA 887

RESULT 14
 AY058173 2017 bp mRNA linear PLN 04-NOV-2001
 Arabidopsis thaliana At1g31800/68069_m00159 mRNA, complete cds.

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
 Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis
 Unpublished
 2 (bases 1 to 1017)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
 Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission
 Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGE (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Chen, R., Chen, H.,
 Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
 Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
 Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES

source
 1..2017
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAFL09-10-L12 (R14026)"
 /ecotype="Columbia"
 1..38
 39..1826
 /note="putative cytochrome P450"
 /codon_start=1
 /product="At1g31800/68069_m00159"
 /protein_id="RA125587.1"
 /db_xref="GI:16648793"
 /translation="MAMAPPLSYTPTITVKPVYSSRRNFVVFSSSSNGRDPLENSV
 PNGVKSLEKQEEKRAELSAISAGATVAKSPSTVKNGLSKIGIPSNVDFPFD
 WTGSDQDPKVPKAGSIOAVRNEAFFIPLYELFLTYGGIFRLTGPKSLFVSDPSI
 AKHILKONAKYSGITAEILDVFMGKGLIPADGEIWRERRRAIPALHOKYVAAMIS
 LFGASDRLCQKDAALAKBEVEMESLFSRLTLDIGKAVENYDFDITCKRWVEA
 VYTVLREADRSVSPIPWKDIPSRKQVATSLKLDNDTLDLIATCKRWVEA
 BELQPHRYMYNERDPSILHLLASGDVSKQLRDLMTMLIAGHETSNAVLTTWTFYL
 LITFESVVKLQEEVDVIGDRPFIQDMKKLKYTRYVNESLRLYPPQPVLRISID
 NDILGEYPIKKEGEDIPISVNHLRSPHLDHDAEKFNPERWPLDGNPNETNQNFSLP
 FGKPRKICIDMFASFENVVAIMLIRFNFQIAFGAPPVKMTTGATHTTBEGLKLV
 TKRTKPLDIPSPILPMDTSRDEVSSALS"
 1827..2017

3'UTR

Query Match 3.7%; Score 91; DB 15; Length 1017;
 Best Local Similarity 60.9%; Pred. No. 8.8e-11;
 Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGATTCTGCAATGTGCAGAGAGGCTTGTGAGAGTTGCAGCCTTATGC 760
 DB 683 TATGATTATTTATTCGGAGAGCTTCAGATAGGCTTTGTTCAGAGCTTGATGCTGCTC 742

QY 761 AGAAGACGGAAGTGTGTAATATGAAGCGAAGTCTCTCAGATGACACTTGTATCAT 820
 DB 743 ATTGAAGGCGGAGAGTAGAGATGATCACTCTCTCTGTTGACACTTGATATTAT 802

QY 821 TGGGTGTCTCTTTTAACTACAATTCGATCTTTTGACTACTGATGCTGCTCATGCA 880
 DB 803 TGGCAAGGCGGTTTTCAATTACGACTTTGACTCCCTTACTAATGATACCGGTGATCGA 862

QY 881 AGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTGTCTTACTGATCTTCTGCCATATG 940
 DB 863 GCGAGTGACACTGTTCTTAAGAGAGCTGGAAGAGAGTGTTCACCTATTCCTGTTG 922

QY 941 GAA 943
 DB 923 GGA 925

RESULT 15

AY056446 2057 bp mRNA linear PLN 03-OCT-2001
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Arabidopsis thaliana At1g31800/68069_m00159 mRNA, complete cds.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2057)

REFERENCE
AUTHORS
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cdna clones

Unpublished

2 (bases 1 to 2057)

REFERENCE
AUTHORS
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (10-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

TITLE
JOURNAL
COMMENT

ORIGIN	/replace=""			
Query Match	3.7%	Score 91;	DB 15;	Length 2057;
Best Local Similarity	60.9%	Pred. No. 8.8e-11;		
Matches 148;	Conservative 0;	Mismatches 95;	Indels 0;	Gaps 0;
QY	701	TGTGAGAGAGCTATTCTGCAAAATGTGCAGAGAGGCTTGTTCAGAAAGTTGCAGCCCTTATGC	760	
DB	678	TATGATTAGTTTATTCGGAGAAGCTTCAGATAGGCTTGTTCAGAAAGTTGTGATGCTGCTGC	737	
QY	761	AGAAGACGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGTCAT	820	
DB	738	ATTGAAAGGGGAAGAAGTAGAGATGGAATCACTCTCTCTCGTTTGACACTTGATATTAT	797	
QY	821	TGGGTGTCTCTTTTAACTACAATTTCGATTCCTTGTGACTACTGATAGTCCCTGTCTATGA	880	
DB	798	TGGCAAGCGGTTTTCAATTACGACTTTGACTCCCTTACTAATGATACCGGTGTGATCGA	857	
QY	881	AGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTTCTGCCCATATTTG	940	
DB	858	GGCAGTGACACTGTTCTTAAGAGAAGCTGAAGACAGAAAGTGTTCACCTATTTCCTGTTTG	917	
QY	941	GAA 943		
DB	918	GGA 920		
Search completed: January 13, 2006, 04:30:26				
Job time : 12354 secs				

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

Location/Qualifiers

1. .2057

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosomes="1"

/clone="RAFL09-33-D23 (R14331)"

/ecotype="Columbia"

1. .33

34. .1821

/notes="putative cytochrome P450"

/codon_start=1

/product="Atlg31800/68069_m00159"

/protein_id="AAL08302.1"

/db_xref="GI:15912337"

/translation="MAKAPLSVTPITVKPVTSRSPNVFSSSSNGRDPLENSV
PWGSKLEKQEKRAELSAIASGAFTRKSGSPSTVNGLSKIGIPSNVLDPMFD
WTGSDQDPKVPKPAKSIQAVNEAFFIPLYELFYGGIFRITFGPKSFLIVSDPSI
AKHILKDNAKYSGLTAEILDFVMGKLI PADGEIWRERRRAIVPALHOKYVAMIS
LFGEASDRCLKDAALKEEVESELSRLTLDIIGKAVNFYDFSLTNDTGVIEA
VTVLREADRSVPPIVWDIPWKDISPRQKVATSLKLIINDLIDLATCKRMVEE
BEIQFHEEYNERDPSILHPLLASGDDVSKOLRDLMTWLIAGHETSAVLTWTFYL
LITPESVVAKLQEVDSVIGDRPTIQDMKKLKYTRVNESLRLYPQPPVLRSLD
NDILGYPIKRGEDIFISVNLHRSPLHWDIAEFNRPWPLDGNPNETNQNFSLP
FGGPRKCIQDMFASPVVAIAMLIIRFNFIAPGAPPVKKMTGTGATITHTTEGLKLT
TKRTPLDIPSPVILPMDTSRDEVSSALS"

1822. .2057

3'UTR

misc_difference 1989

/notes="compared with genome"

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2006, 00:04:14 ; Search time 434 Seconds
(Without alignments)
10104.250 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagttctactcttttc.....tgaaggtgagccaaggtaa 2467

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.2	3.0	1831	3	US-08-948-564-15
2	57.8	2.3	404	3	US-09-615-192A-156
3	57.8	2.3	404	3	US-09-169-789-156
4	49.8	2.0	1141	3	US-09-806-7088-22
5	49.2	2.0	1141	3	US-09-806-7088-22
6	48.8	2.0	832	3	US-09-621-976-2813
7	46.8	1.9	7218	2	US-08-232-463-14
8	44.6	1.8	1664976	3	US-08-916-421B-1
9	44.6	1.8	1664976	3	US-09-692-570-1
10	41.8	1.7	43192	3	US-09-949-016-15466
11	41.4	1.7	601	3	US-09-949-016-130943
12	41.4	1.7	67899	3	US-09-949-016-15432
13	41.4	1.7	67902	3	US-09-949-016-11870
14	40	1.6	832	3	US-09-621-976-2813
15	39.6	1.6	601	3	US-09-949-016-88266
16	39.6	1.6	601	3	US-09-949-016-88267
17	39.6	1.6	49011	3	US-09-949-016-14221
18	39.6	1.6	168394	3	US-09-949-016-13002
19	39.4	1.6	1125	3	US-09-248-796A-6024
20	39.4	1.6	1308	3	US-09-902-540-3557
21	39.4	1.6	19019	3	US-09-902-540-1171
22	39	1.6	672	3	US-09-248-796A-1435
23	38.8	1.6	58407	3	US-08-916-421B-2
24	38.8	1.6	58407	3	US-09-692-570-2

c	25	38.2	1.5	7218	2	US-08-232-463-14	Sequence 14, Appl
	26	38.2	1.5	135476	3	US-09-949-016-12611	Sequence 12611, A
	27	38.2	1.5	135476	3	US-09-949-016-14413	Sequence 14413, A
c	28	38	1.5	601	3	US-09-949-016-200471	Sequence 200471, A
c	29	38	1.5	32392	3	US-09-662-2548-27	Sequence 27, Appl
	30	37.2	1.5	113966	3	US-09-949-016-12277	Sequence 12277, A
	31	37.2	1.5	113967	3	US-09-949-016-17051	Sequence 17051, A
	32	37	1.5	268449	3	US-09-949-016-17244	Sequence 17244, A
	33	36.8	1.5	2078	3	US-09-949-016-2102	Sequence 2102, Ap
	34	36.8	1.5	3150	3	US-10-018-730A-3	Sequence 3, Appl
c	35	36.8	1.5	4312	3	US-10-104-047-468	Sequence 468, App
c	36	36.8	1.5	74177	3	US-09-949-016-11988	Sequence 11988, A
c	37	36.8	1.5	74177	3	US-09-949-016-17388	Sequence 17388, A
	38	36.6	1.5	1584	3	US-09-351-223-1	Sequence 1, Appl
	39	36.4	1.5	601	3	US-09-949-016-92564	Sequence 92564, A
	40	36.4	1.5	601	3	US-09-949-016-92565	Sequence 92565, A
	41	36.4	1.5	1473	3	US-08-277-031B-11	Sequence 11, Appl
	42	36.4	1.5	1473	3	US-08-277-031B-12	Sequence 12, Appl
	43	36.4	1.5	1473	3	US-08-277-031B-13	Sequence 13, Appl
	44	36.4	1.5	1829	2	US-08-201-118-8	Sequence 8, Appl
	45	36.4	1.5	1829	2	US-08-238-821B-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-08-948-564-15

; Sequence 15, Application US/08948564

; Patent No. 6121512

; GENERAL INFORMATION:

; APPLICANT: Siminszky, Balazs

; APPLICANT: Dewey, Ralph E.

; APPLICANT: Corbin, Frederick T.

; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and Methods of Producing Herbicide-Resistant Transgenic Plants

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Virginia C. Bennett

; STREET: PO Box 37428

; CITY: Raleigh

; STATE: No. 6121512th Carolina

; COUNTRY: USA

; ZIP: 27627

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/948,564

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Bennett, Virginia C.

; REGISTRATION NUMBER: 37,092

; REFERENCE/DOCKET NUMBER: 5051-409

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-854-1400

; TELEFAX: 919-854-1401

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1831 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 20..1747

US-08-948-564-15

Query Match 3.0%; Score 73.2; DB 3; Length 1831;


```
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match          2.0%; Score 48.8; DB 3; Length 832;
Best Local Similarity 10.4%; Pred. No. 0.0012;
Matches 35; Conservative 163; Mismatches 140; Indels 0; Gaps 0;

QY 1897 TTCATGATCTTTGTGTGGTTCGGATATCTAACCGAGTGGACATTCCTAGTATTACATT 1956
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 WKYTTWYAKCWTKKWSNSYMMYKWKYTWRRWRKKKCAWKYKTKTWYWRWYAMW 65

QY 1957 CATGCCACATTTCTATGTGTGTTGTTGTTGTTTATTCAAAGTGATGGGAAAAGCTGA 2016
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 GYKKKACMCTRTKKKKKKGYMMYWGWRWSYMAWTRTWGYAYRSMYTWWRRCWK 125

QY 2017 GCAATTTCTGCTGACGAGTTCGACATAGTCGCGCAATCCCTAACGAAACAAACACTGA 2076
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 KKAYRKTTCYSSKGTWWRKKKATTTWKKTYWATRYWMMWTKWRBSWYCW 185

QY 2077 TTTCAAGTAACTCAGTAGACACATCTTTTGACACAAACTACTGAATCAAGATTAGTG 2136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 WNGKARKWSTWRSKRSYASARSACRCCYSCSWGMSWKMWRWWRGWTAGMKAWRA 245

QY 2137 TTTTGATYAGGAATTTAAAGATGATTTCTTTTTCACAGATTCATCCCATTCAGTG 2196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 SCWRRKYAGSKTSYKSMWMCWTRSWKYCYTKARWTGYCYRKGGMWGRGRWYASKY 305

QY 2197 GAGGCGCTAGAAAATCTGTAGCGATCAGTTTGCAATTG 2234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 MWKRWWCWARMTRISGTGRASWWRWRYTMMKWK 343

RESULT 7
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match          1.9%; Score 46.8; DB 2; Length 7218;
Best Local Similarity 0.8%; Pred. No. 0.015;
Matches 3; Conservative 215; Mismatches 142; Indels 0; Gaps 0;

QY 7 TCTTCACATCTTTCTCCATCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1078 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1137

QY 67 CGTCTTTTATCACCAAAACCCAAATTCATTTCTCATCATCATCTCTCATTCAGAGAACCC 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1138 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1197

QY 127 AAACCAACTCGAGACCAATTCATCGAAATCCCAATCATGGTCAGTCCCGATTGGCTC 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1198 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1257

QY 187 ACAACACTCAGTCGACCTTTCTCTCAGGAAAAACGACGATCAGGTATACCAATCGCG 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1258 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1317

QY 247 AACGGAAGCTCGACGATGTCGCTCATCTCTCGGAGGTGCTCTCTCTCTCTCTCTCTAC 306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1318 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1377

QY 307 AAATGATGAATGAGTACGAGCCCATTTACCGTCTCGCTGCTGCTGCTGCTGCTGCTGTA 366
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1378 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1437

RESULT 8
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; Patent No. 6503729
; FILE REFERENCE: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
```



```

; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1569020)..(1569020)

```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match      1.8%; Score 44.6; DB 3; Length 1664976;
Best Local Similarity 50.2%; Pred. No. 1.4;
Matches 110; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1060 GAGACAGGTGAAGCTGAAAGCGTGTAACTTTGATAGGGAACCTGTTGAGACCTTAT 1119
Db 1435967 GAGAGAAATGACATCAAGAGGAATTAATGGAATTAATTGAAGCAATTAATATACGTC 1436026
QY 1120 TCCTAAGTCTAAGAAATTTGCGAAGAGAGCGCAAGAAATCAATGATGAGGAGTATGT 1179
Db 1436027 TGAAGAACTGGAAGAGATTTGTCATGGTAAAGACCAAGATCATTTTAAAGAGAGTAG 1436086
QY 1180 AATGATGTCGCCCAAGTATCCTCGTTTTCTTGTGCAAGCAGAGAGAGGTTTAAAC 1239
Db 1436087 GATTATTGATGTTCAAGGAGATGAAGGAATATATTAGAAGGAAGGAGGATGGAAA 1436146
QY 1240 TTTTTCCTTAAGTTTATTAAGCAAAATTTGGCCCTTTTCATT 1278
Db 1436147 GATAAAGGCAAGATTAATTGTTTAAAAAAGGCTATAAATT 1436185

RESULT 9
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
```

```
;
;
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319236)..(319236)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1095846)..(1095846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
```

```
;
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
;
;
; Query Match 1.8%; Score 44.6; DB 3; Length 1664976;
; Best Local Similarity 50.2%; Pred. No. 1.4;
; Matches 110; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
;
; QY 1060 GAGACAGGTGAAAGCTGAAAAGCGCTGTAACCTTTGATAAGGGAACCTGTTGAAGACCTTAT 1119
; DB 1435967 GAGAGAAATGAGCATCAAAAGAGGAAATTAATGGAATTAATTGAAGCAATTAATATATACGTC 1436026
;
; QY 1120 TCGTAAGTGTAAAGAAATGTCGAAGAGAGAGCGGAAGCAATCAATGATGAGGAGTATGT 1179
; DB 1436027 TGAATAAACCTGAAGAGATGTTTCATGTTAAGGACCAAGATCATTGTTAAGAGAGTAG 1436086
;
; QY 1180 AAATGATGCTGACCCCAAGTATCCTGCGTTTCTTGTTCGACGAGAGAGAGGTTTAAAC 1239
; DB 1436087 GATTATTGATGTTCAAGGAGATGAAGGATATATATTAGAAGGGAAGAGAGGATGAAA 1436146
;
; QY 1240 TTTTTCCTTAAGTTTATAGCAAAATTTGCGCTTTTCATT 1278
; DB 1436147 GATAAAGGCAAGATTAATTGTTTAAAAAGGCTATAAATT 1436185
;
; RESULT 10
; US-09-949-016-15466
; Sequence 15466, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15466
; LENGTH: 43192
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15466
;
; Query Match 1.7%; Score 41.8; DB 3; Length 43192;
; Best Local Similarity 59.9%; Pred. No. 1.2;
; Matches 88; Conservative 0; Mismatches 57; Indels 2; Gaps 1;
;
; QY 770 AAGTCTGTGATATGGAAGCGAAGTCTCTCAGATGACACTTGATGATGGGTGTC 829
; DB 1863 AAGTCTGGAGTATAGCGGTGAGCCACGTCCTGGCCCTAGATTCCATTGTTATCTC 1922
```

```
QY 830 TCTTTTAACTACAA--TTTCGATCTTTGACTACTGATAGTCCCTGTCATTTGAAGCTGTT 887
Db 1923 TGTCTTAAAGCAAAACTTTTTCCTCTTAGACTTCAGATTTTTCCTGCTTGATCCCGAC 1982
QY 888 TACACTGCTCTTAAAGAAAGCTCAGCTT 914
Db 1983 TGGACTGCTCAAGAAGAAATGSCCTT 2009

RESULT 11
US-09-949-016-130943/c
; Sequence 130943, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130943
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-130943

Query Match 1.7%; Score 41.4; DB 3; Length 601;
Best Local Similarity 56.1%; Pred. No. 0.14; Mismatches 0; Indels 0; Gaps 0;
Matches 78; Conservative 0;

QY 1044 TGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACTTTTGATAAGGGAAA 1103
Db 253 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAAATTCCTTAGTAGGATAA 194
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 193 AGTTTGATATCTTCAGTACAAATTTGTACAACATAATTTTCATAAATAAGGAAAAGGCTTA 134
QY 1164 ATGATGAGGAGTATGTAAA 1182
Db 133 ATTTTCAGGTAGACTGAAA 115

RESULT 12
US-09-949-016-15432/c
; Sequence 15432, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15432
; LENGTH: 67899

QY 1044 TGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACTTTTGATAAGGGAAA 1103
Db 26889 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAAATTCCTTAGTAGGATAA 26830
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 26829 AGTTTGATATCTTCAGTACAAATTTGTACAACATAATTTTCATAAATAAGGAAAAGGCTTA 26770
QY 1164 ATGATGAGGAGTATGTAAA 1182
Db 26769 ATTTTCAGGTAGACTGAAA 26751

RESULT 13
US-09-949-016-11870/c
; Sequence 11870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11870
; LENGTH: 67902
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11870

Query Match 1.7%; Score 41.4; DB 3; Length 67902;
Best Local Similarity 56.1%; Pred. No. 2; Mismatches 0; Indels 0; Gaps 0;
Matches 78; Conservative 0;

QY 1044 TGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACTTTTGATAAGGGAAA 1103
Db 26889 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAAATTCCTTAGTAGGATAA 26830
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 26829 AGTTTGATATCTTCAGTACAAATTTGTACAACATAATTTTCATAAATAAGGAAAAGGCTTA 26770
QY 1164 ATGATGAGGAGTATGTAAA 1182
Db 26769 ATTTTCAGGTAGACTGAAA 26751

RESULT 14
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
```

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15432

Query Match 1.7%; Score 41.4; DB 3; Length 67899;
Best Local Similarity 56.1%; Pred. No. 2; Mismatches 0; Indels 0; Gaps 0;
Matches 78; Conservative 0;

QY 1044 TGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACTTTTGATAAGGGAAA 1103
Db 26889 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAAATTCCTTAGTAGGATAA 26830
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 26829 AGTTTGATATCTTCAGTACAAATTTGTACAACATAATTTTCATAAATAAGGAAAAGGCTTA 26770
QY 1164 ATGATGAGGAGTATGTAAA 1182
Db 26769 ATTTTCAGGTAGACTGAAA 26751
```

```

; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

```

Query Match 1.6%; Score 40; DB 3; Length 832;

	Best, Local Similarity	8.6%;	Pred. No. 0.44;	Mismatches 159;	Indels 139;	Gaps 0;
1904	ATCTTTGTGTTGGGATATCTAAACGGAGTGGACATTCCTAGTATTACATTCATGCC	1963	1904	ATCTTTGTGTTGGGATATCTAAACGGAGTGGACATTCCTAGTATTACATTCATGCC	1963	
329	WKSTVACASRYKRYKWTGWWTMMKRMWSTWYCYMKCKMYRGRRCAWYTMARGRWWSY	270	329	WKSTVACASRYKRYKWTGWWTMMKRMWSTWYCYMKCKMYRGRRCAWYTMARGRWWSY	270	
1964	ACATTTCTTAGTGTGTTGTTGTTATTCTCAAGGTATCGGGAAGAAGCTGAGGAATTT	2023	1964	ACATTTCTTAGTGTGTTGTTGTTATTCTCAAGGTATCGGGAAGAAGCTGAGGAATTT	2023	
269	AWGKSKMRSAMSMTCTMYKYYKSGSYTMTKTCATWYWKYWKRMWKTCSGSRGGY	210	269	AWGKSKMRSAMSMTCTMYKYYKSGSYTMTKTCATWYWKYWKRMWKTCSGSRGGY	210	
2024	CTGCCTGAACGATTCGACATAGATGGCGCAATCCCTAAACGAAACAAACACTGATTTCAAG	2083	2024	CTGCCTGAACGATTCGACATAGATGGCGCAATCCCTAAACGAAACAAACACTGATTTCAAG	2083	
209	MTSYTSTESYSMYASWMYTWCWNGRWWSYTWMAWGKKWRVATTTWRRAAMWAAWT	150	209	MTSYTSTESYSMYASWMYTWCWNGRWWSYTWMAWGKKWRVATTTWRRAAMWAAWT	150	
2084	TAAACTCAGTAGAACACATCTTTTGTGACACAAACTACTGAATCAAGATTAGTGTGTTTGTAT	2143	2084	TAAACTCAGTAGAACACATCTTTTGTGACACAAACTACTGAATCAAGATTAGTGTGTTTGTAT	2143	
149	MMWYMMWAAWCHSSRGAAYRRTMMWGYRYWRKCKSYRTRCAWAYAKWTKRSYTWCVRW	90	149	MMWYMMWAAWCHSSRGAAYRRTMMWGYRYWRKCKSYRTRCAWAYAKWTKRSYTWCVRW	90	
2144	TAGGGAATTTAAAGAAGATGATTTCTTTTTTCCACAGATTCATCCATCTCAGTGGAGGCC	2203	2144	TAGGGAATTTAAAGAAGATGATTTCTTTTTTCCACAGATTCATCCATCTCAGTGGAGGCC	2203	
89	KWKCKMMWMAAYGKTMWRBAWCKTYRWKRWAAWWRMWTMMWYTWTRAKCRWM	30	89	KWKCKMMWMAAYGKTMWRBAWCKTYRWKRWAAWWRMWTMMWYTWTRAKCRWM	30	
2204	TAGAAATATGTCAGCGCATCAGTTTG	2229	2204	TAGAAATATGTCAGCGCATCAGTTTG	2229	
29	WRKWRSSWMMWMAWGMTRWAAARMWR	4	29	WRKWRSSWMMWMAWGMTRWAAARMWR	4	

RESULT 15

```

US-09-949-016-88266/c
; Sequence 88266, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR APPLICATION NUMBER: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR APPLICATION NUMBER: 2000-10-03
; PRIOR FILING DATE: 60/231,498
; PRIOR APPLICATION NUMBER: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 88266
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88266

```

Query Match
1.6%; Score 39.6; DB 3; Length 601;

Best Local Similarity 49.1%; Pred. No. 0.48;
Matches 105; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

791 GAAGTCTCTCAGATGACACTGTGATGTCATTGGGTGTCTCTTTTAACTACAAATTTTGA 850

[illegible]

Search completed: January 13, 2006, 07:19:43
Job time : 438 secs

This Page Blank (uspto)

		Matches 2466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1801	TTATGATTCAGTCTATTAACATCCATCTTCTCGAGGTACAGTTCCTTCCTCTCTC	1860
Db	2810	TTATGATTCAGTCTATTAACATCCATCTTCTCGAGGTACAGTTCCTTCCTCTCTC	2869
Qy	1861	GTCCATAGTATAACATAGGGAGCTATCTCTCTCAATGATCTTTGTGTGTTGCGG	1920
Db	2870	GTCCATAGTATAACATAGGGAGCTATCTCTCTCAATGATCTTTGTGTGTTGCGG	2929
Qy	1921	ATATCTAACCGAGTGGACATTCCTAGTATTACATTCATGCCCCACATTTCTTATGTGTTT	1980
Db	2930	ATATCTAACCGAGTGGACATTCCTAGTATTACATTCATGCCCCACATTTCTTATGTGTTT	2989
Qy	1981	GTGTGTTGTTATTCCTCAAGGTATGGAAAAAGCTGAGGAAATTCGCTGAAACGATTCGA	2040
Db	2990	GTGTGTTGTTATTCCTCAAGGTATGGAAAAAGCTGAGGAAATTCGCTGAAACGATTCGA	3049
Qy	2041	CATAGATGGCGCAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	2100
Db	3050	CATAGATGGCGCAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	3109
Qy	2101	ATCTTTTGACACAAACTACTGAACTCAAGATTAGTGGTTTGAATAGGGAAATTTAAAGAT	2160
Db	3110	ATCTTTTGACACAAACTACTGAACTCAAGATTAGTGGTTTGAATAGGGAAATTTAAAGAT	3169
Qy	2161	GATTTCTTTTTCACAGATTCATCCATTCAGTGGAGGGCTAGAAAAATGTGTAGGGG	2220
Db	3170	GATTTCTTTTTCACAGATTCATCCATTCAGTGGAGGGCTAGAAAAATGTGTAGGGG	3229
Qy	2221	ATCAGTTGTCATTGATGAGGCAATTTGGCACTCGCGGTGTTCTTCAGCGGTTAAAGG	2280
Db	3230	ATCAGTTGTCATTGATGAGGCAATTTGGCACTCGCGGTGTTCTTCAGCGGTTAAAGG	3289
Qy	2281	TTGAGTGTGTTCTGATCAGACCATTTAGCATGACCAACAGGAGCAACCATACACACCA	2340
Db	3290	TTGAGTGTGTTCTGATCAGACCATTTAGCATGACCAACAGGAGCAACCATACACACCA	3349
Qy	2341	ATGTATGCCAATGTTCTCACCTCGAGAGATTAATGAGAGTGTCTGTTTGTAGATG	2400
Db	3350	ATGTATGCCAATGTTCTCACCTCGAGAGATTAATGAGAGTGTCTGTTTGTAGATG	3409
Qy	2401	ATTCGAATTTCTAATGCTGATATTTTCAATTTGAGGATTTGATATGAGGTGAGCCA	2460
Db	3410	ATTCGAATTTCTAATGCTGATATTTTCAATTTGAGGATTTGATATGAGGTGAGCCA	3469
Qy	2461	AAAGTAA 2467	
Db	3470	AAAGTAA 3476	
RESULT 3			
US-10-751-235-7			
; Sequence 7, Application US/10751235			
; Publication No. US20050150002A1			
; GENERAL INFORMATION:			
; APPLICANT: DellaPenna, Dean			
; APPLICANT: Tian, Li			
; APPLICANT: Kim, Joonvul			
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid			
; FILE REFERENCE: MSU-08604			
; CURRENT APPLICATION NUMBER: US/10/751,235			
; CURRENT FILING DATE: 2004-01-02			
; NUMBER OF SEQ ID NOS: 74			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 7			
; LENGTH: 2467			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-10-751-235-7			
Query Match 99.9%; Score 2465.4; DB 9; Length 2467;			
Best Local Similarity 100.0%; Pred. No. 0;			


```
QY 361 TTGCTAAATTGTGAGCGACCCAGCGATAGCTAAACATGTTTGGAGAAATTATCCAAAGTAC 420
DB |||||||
QY 439 TTGCTAAATTGTGAGCGACCCAGCGATAGCTAAACATGTTTGGAGAAATTATCCAAAGTAC 498
DB |||||||
QY 421 GCTAAAGGCTTAGTGGCTGAACTCTCTGAAATTTCTTAATTTGGTTCGGGTTTCGCTATCGCT 480
DB |||||||
QY 499 GCTAAAGGCTTAGTGGCTGAACTCTCTGAAATTTCTTAATTTGGTTCGGGTTTCGCTATCGCT 558
DB |||||||
QY 481 GAAGACCTCTTTTGGACAGTAATTTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB |||||||
QY 559 GAAGACCTCTTTTGGACAGTAATTTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
DB |||||||
QY 541 TTGTGGAGTAATGTGTGACTGTCTTGTATGATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 600
DB |||||||
QY 577 ----- 576
DB |||||||
QY 601 TCCAATCTTCTCTATTTGGGCTTAGCTGAAGTCTGATTTTTTATAGCGGAGCGTAGAG 660
DB |||||||
QY 577 ----- 589
DB |||||||
QY 661 CGGTGGTTCCATCGCTTACAGAGGATATTTGTCTGTGATTTGTGTGAGAGAGTATTTCTGCA 720
DB |||||||
QY 590 CGGTGGTTCCATCGCTTACAGAGGATATTTGTCTGTGATTTGTGTGAGAGAGTATTTCTGCA 649
DB |||||||
QY 721 AATGTGCAGAGAGGCTTTGTGAGAGTTCAGCCCTTATGCAAGACGGAAGTGTCTGCA 780
DB |||||||
QY 650 AATGTGCAGAGAGGCTTTGTGAGAGTTCAGCCCTTATGCAAGACGGAAGTGTCTGCA 709
DB |||||||
QY 781 ATATGGAAGCAAGTCTCTCAGATGACACTTGTGATCTCAATGGGTTGTCTCTTTTAACT 840
DB |||||||
QY 710 ATATGGAAGCAAGTCTCTCAGATGACACTTGTGATCTCAATGGGTTGTCTCTTTTAACT 769
DB |||||||
QY 841 ACAATTTGCAATCTTTGACTACTGATAGTCTGTCTGATGAGTGTGATGAGTGTGATGAGTGTGAT 900
DB |||||||
QY 770 ACAATTTGCAATCTTTGACTACTGATAGTCTGTCTGATGAGTGTGATGAGTGTGATGAGTGTGAT 829
DB |||||||
QY 901 AAGAGCTGAGCTTGGTCTTACTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB |||||||
QY 830 AAGAGCTGAGCTTGGTCTTACTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871
DB |||||||
QY 961 TTTTCTGTGGTGTGTTGATTTGTGTGGAACAAATGGATTTCTTGTAAATGAGAGGGTGTGG 1020
DB |||||||
QY 872 ----- 871
DB |||||||
QY 1021 TTGTTTTTTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAGGTGAAGCTGAAA 1080
DB |||||||
QY 872 -----AGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAGGTGAAGCTGAAA 920
DB |||||||
QY 1081 GGCTGTAACTTTGATAGGGAACCTGTTGAGACCTTATTGCTTAAGTGTAAAGAAATCT 1140
DB |||||||
QY 921 GGCTGTAACTTTGATAGGGAACCTGTTGAGACCTTATTGCTTAAGTGTAAAGAAATCT 980
DB |||||||
QY 1141 CGAAGAGAGGCGAAGGAATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 1200
DB |||||||
QY 981 CGAAGAGAGGCGAAGGAATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 1040
DB |||||||
QY 1201 CTTGCGTTTCTTGTCTGCAAGCAGAGAGAGGTTTAAACCTTTTCTTAAAGTTTAAAG 1260
DB |||||||
QY 1041 CTTGCGTTTCTTGTCTGCAAGCAGAGAGGTTTAAACCTTTTCTTAAAGTTTAAAG 1069
DB |||||||
QY 1261 CAATTTGGGCTTTCAATTCGCAATTCGAGCTGATGTTGCAATTTGAGGTTTTCAG 1320
DB |||||||
QY 1070 -----AG 1071
DB |||||||
QY 1321 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGGGGTCAATGAACC 1380
DB |||||||
QY 1072 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGGGGTCAATGAACC 1131
DB |||||||
QY 1381 ACTGGATCTGTCTCTCACTGGACACTTTTATCTCTCAAGTAAGTGA CTTTAAATGTATCTTC 1440
DB |||||||
QY 1132 ACTGGATCTGTCTCTCACTGGACACTTTTATCTCTCAAGTA----- 1171
DB |||||||
QY 1441 TACTTTGCTATGTAGAGAAATTTACTTGGATGGAGGCTTCTCTGTTCTCATTTACCTCTT 1500
DB |||||||
```

```
DB 1172 ----- 1171
QY 1501 CAATTTCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAAGCAACAAGAAAGTAGACA 1560
DB |||||||
QY 1172 -----AGAACTCATCTGCAATTAAGGAAAGCAACAAGAAAGTAGACA 1213
DB |||||||
QY 1561 GAGTGTAGAAAGGAAAGAAACCGGCTTTCGAGGATATAAAGGAGTTGAAGTACATCATCTC 1620
DB |||||||
QY 1214 GAGTGTAGAAAGGAAAGAAACCGGCTTTCGAGGATATAAAGGAGTTGAAGTACATCATCTC 1273
DB |||||||
QY 1621 GTTGTATAAAACGAGTCAATGCGTCTATCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
DB |||||||
QY 1274 GTTGTATAAAACGAGTCAATGCGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1314
DB |||||||
QY 1681 TCTCTAATTTTCATGAACATAAATTTTCTGATTTGTTTCTCTGTTAGGTTCTTGATAA 1740
DB |||||||
QY 1315 -----GTCCTTGATAA 1324
DB |||||||
QY 1741 GAAGAGCTCAAGTTCCTGACATTTCTTCTGGGAACTATAAGGTCAATACCGGACAAAGCA 1800
DB |||||||
QY 1325 GAAGAGCTCAAGTTCCTGACATTTCTTCTGGGAACTATAAGGTCAATACCGGACAAAGCA 1384
DB |||||||
QY 1801 TTATGATTTTCAAGTCTTATAACATCCATCGTCTTCTCCGAGGTA 1841
DB |||||||
QY 1385 TTATGATTTTCAAGTCTTATAACATCCATCGTCTTCTCCGAGGTA 1425
DB |||||||
```

RESULT 5

```
US-10-425-114-32608
; Sequence 32608, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32608
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17009E12_FLI
US-10-425-114-32608
```

```
Query Match 7.9%; Score 195.8; DB 7; Length 1804;
Best Local Similarity 63.2%; Pred. No. 4.3e-42;
Matches 371; Conservative 0; Mismatches 127; Indels 89; Gaps 1;
```

```
QY 647 GCGAGGCGGTAGAGCGGTGTTCCATCGCTTACAGGAGGTATTTCTGTGATTTCTGGA 706
DB |||||||
QY 394 GGTGAGACGACAGAGCTGTTGTGCTTCCCTGCAAAAAGATTTCTGTCCATTAATCGTGA 453
DB |||||||
QY 707 GAGAGTATTCTCAAAATGTCAGAGAGGCTTGTGAGAAAGTTGACGCCCTTATGCAAGA 766
DB |||||||
QY 454 CRAAGTATTTTGTAAATGCTGTGAGAGACTAATAGACAAACTTGAGCCATATGCTTTGAG 513
DB |||||||
QY 767 CGGAGTGTGTGAATATGAAAGCGAAGTTCTCTCAGATGACACTTGTGATGCTTGGGTT 826
DB |||||||
QY 514 TCGGGAACTGTCAATATGGAAGCGAGGTTTCTCAGTTGACATTCGATTCGATTCGTTT 573
DB |||||||
QY 827 GTCTTTTAACTACAATTTGATTTCTTGTACTACTGATGATCTGCTCATTCGAAGCTGT 886
DB |||||||
QY 574 ATCATTTGTTCAACTACAAATTTGATTTCCCTCAACAGATAGTCTCTGTCATTGATCTGT 633
DB |||||||
```

QY 887 TTACTGCTCTTAAGAGCTGAGCTGGTTCTTACTGATCTTCTGCCATATTTGAAGGC 946
Db 634 TTATACTGCACCTCAAGAGCAGAGCTTGGTTCTACAGATCTTTTGGCCACTGGAAGGT 693
QY 947 AAGTTTCTCTGCTGTTTTTCTGGTTTGTGATTGTGTGGAACAATTGGATTCTTTGTTAA 1006
Db 694 TGGTTTC----- 700
QY 1007 TTGAGAGGTTTGGTTTCTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
Db 701 -----TTGTGCAAGATAATCCCAAGACAG 724
QY 1067 GTGAAGCTGAAAGCTGTAACCTTTGATAAGGGAACCTTTGAGACCTTATTCCTAAG 1126
Db 725 ATAAAGCAGAGAATGCGGTTACGATTATAAGGAACACTGTTGAAGAGCTGATTATGAAG 784
QY 1127 TGTAAAGAAATTTGTCCAAAGAGAGCGGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
Db 785 TGTAAAGAAATAGTGGAGCTGAAATGAACAGATTGAGGTTGAGGAATATGTAACGNA 844
QY 1187 GCTGACCAAGTATCTCGGTTTCTTGTGTCGAAGCAGAGAAGAGGT 1233
Db 845 GGGGATCCTAGCATTTCTACGCTTCTCTACTTGTAGCCGAGATGAGGT 891

RESULT 6

US-10-425-115-148190
; Sequence 148190, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 148190
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_66678C.1
US-10-425-115-148190

Query Match 7.9%; Score 195.8; DB 8; Length 2537;
Best Local Similarity 63.2%; Pred. No. 5.3e-42;
Matches 371; Conservative 0; Mismatches 127; Indels 89; Gaps 1;
QY 647 GCGAGGCGTAGAGCGGTGTTCCATCGCTTCACAGGAGGTATTTGTCTGTGATTGTGGA 706
Db 647 GGTGAGCGCAGAGCTGTTGGCTTCCCTCGCAAAAGATTCTGTCCATATTCGTGGA 706
QY 707 GAGAGTATTTGCAAAATGTCAGAGAGCGCTTGTGAGAAAGTTGCAAGCCTTATGCAAGA 766
Db 707 CAAGTATTTTGTAAATGCTGTGAGAGACTAATAGACAAACTGTAGCCATATGCTTTGAG 766
QY 767 CGGAAGTCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTGATGCTGGTT 826
Db 767 TGGGAACTCTGCAATATGGAAGCGAGGTTTCTCAGTTGCAATTTGGATGTTGGTTT 826
QY 827 GTCTCTTTTAACTACAAATTCGATTCTTTGACTACTGATGCTGCTGCTGCTGCTGCTGCT 886
Db 827 ATCATTTGTCAACTACAAATTTGATTTCCCTCACAACAGATAGTCTGCTGCTGCTGCTGCT 886
QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCTGCTTCTACTGATCTTCTGCTGCTGCTGCTGCTGCT 946
Db 887 TTACTGCTCTTAAAGAGCAGAGCTTCTGCTTCTACTGATCTTCTGCTGCTGCTGCTGCTGCT 946
QY 947 AAGTTTCTCTGTTTTTTTCTGTTGTTGTTGATTGTGTGGAACAATTTGGATTCTTTGTTAA 1006

Db 947 TGGTTTC----- 953
QY 1007 TTGAGAGGTTTGGTTTCTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
Db 954 -----TTGTGCAAGATAATCCCAAGACAG 977
QY 1067 GTGAAGCTGAAAGCGTGTAACTTTGATAAGGGAACCTTTGAGAACCTTATTCCTAAG 1126
Db 978 ATAAAGCAGAGAATGCGGTTACGATTATAAGGAACACTGTTGAAGAGCTGATTATGAAG 1037
QY 1127 TGTAAAGAAATTTGTCCAAAGAGAGCGGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
Db 1038 TGTAAAGAAATAGTGAAGAGCTGAAATGAACAGATTGAGGTTGAGGAATATGTAACGAA 1097
QY 1187 GCTGACCAAGTATCTCGGTTTCTTGTGTCGAAGCAGAGAAGAGGT 1233
Db 1098 GGGGATCCTAGCATTTCTACGCTTCTCTACTTGTAGCCGAGATGAGGT 1144

RESULT 7

US-10-751-235-24
; Sequence 24, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-751-235-24

Query Match 7.6%; Score 187.2; DB 9; Length 1086;
Best Local Similarity 61.9%; Pred. No. 7.2e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;
QY 647 GCGAGGCGTAGAGCGGTGTTCCATCGCTTCACAGGAGGTATTTGTCTGTGATTGTGGA 706
Db 489 GGTGAGCGTAGAGCAGTTGTACCATCTCTACACAAAGATTCTCTCAGTAATGCTCGA 548
QY 707 GAGAGTATTTGCAAAATGTCAGAGAGCGCTTGTGAGAAAGTTGCAAGCCTTATGCAAGA 766
Db 549 TAAAGTGTCTGTAATGCTGTGAGAGATTGTTGGAAGGCTCGAGACTTATGCTTTGAG 608
QY 767 CGGAAGTCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTGATGCTGGTT 826
Db 609 TGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAATGACATTAGATGATTGGTTT 668
QY 827 GTCTCTTTTAACTACAAATTTGATTTCTTGACTACTGATAGTCTGCTGCTGCTGCTGCTGCT 886
Db 669 ATCTTGTTCACCTACAACTTTGATTCCCTCACATCAGATAGTCTGTTTATGATGCTGT 728
QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCTGCTTCTACTGATCTTCTGCTGCTGCTGCTGCTGCT 946
Db 729 TTACTGCTCTTAAAGAGCTGAGGCTGCTTCTACAGATCTTTTACCATACTGCTGCTGCTGCT 783
QY 947 AAGTTTCTGTTTTTTTCTGTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTAA 1006
Db 784 ----- 783
QY 1007 TTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
Db 784 -----CAGATCGATTGCTGTGCAAGATTGTTCTTCTAGACAG 819

APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Query Match 7.4%; Score 183; DB 9; Length 1686;
Best Local Similarity 61.8%; Pred. No. 1.3e-38;
Matches 363; Conservative 0; Mismatches 135; Indels 89; Gaps 1;

```
QY 647 GCGAGGCTAGAGCGGTGCTCCATCGCTTACAGGAGGTATTTGCTGTGATGTGCA 706
DB 522 GGTGAGCGTGCATCAGTTGTACCATCTCTACACAACGATTTCTCTCGGTGATGTTGA 581
QY 707 GAGAGTATTCTGCAATGTCAGAGAGGCTTGTGAGAAAGTTGCAGCCTTATGCGAAGA 766
DB 582 CAGAGTTTTTGTAAATGTCTGAGAGATTAGTGGAGAGCTTGAGACATCTGCTTAA 641
QY 767 CGAAGTGTCTGAATATGGAAGCGAAGTTCTCTCAGATGACATTTGATGCTATGGGTT 826
DB 642 TGGCAAACTGTAAATATGGAAGCAAGGTTCTCTCAATGACTTTAGATGCTATGGTTT 701
QY 827 GTCTCTTTTAACTCAATTTCCGATTTCTTTGACTACTGATGCTGCTGCTGCTGCTG 886
DB 702 GTCTCTTTTAACTCAATTTTGAATTCCTCCTCAGATAGCCCTGTTATTTGATGCTGT 761
QY 887 TTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGC 946
DB 762 TTACACTGCACTCAAGAGAGCAACTTCTGTTCTACAGATCTTTTACCACTAGCA---- 817
QY 947 AAGTTTCTGTGTTTTCTGTGTTTTGTTGATTTGTGGAACAATTTGGATTTCTTTTAA 1006
DB 818 ----- 817
QY 1007 TTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGATTTGTGTAGATAGTCCCGAGAG 1066
DB 818 -----AGATTGATTTGCTGTGTCGAAGATTGTTCTCTAGACAA 852
QY 1067 GTGAAGCTGAAAGCGCTTAATCTTTGATAAGCGAACTGTTTGAAGACCTTATTTGTAAG 1126
DB 853 ATAAAGCGAAGAGCGAGTTAAATCATCATCAGAAATACGTTGAGGACCTAATTACCAA 912
QY 1127 TGTAAAGAAATTTGTGAAAGAGAGCGGAAAGAATCAATGATGAGGAGTATGTAATGAT 1186
DB 913 TCGAAGAGATTGTAGTCTGAGAATGAACAAATTTGAGGTTGAGGATATGTAATGAG 972
QY 1187 GTGACCCAGTATCTGCTGCTTCTGCTTTCGCTTGAAGCAGAGAGAGGT 1233
DB 973 CGAGACCTTAGCATCTCGCAATTCCTACTTGTAGCCGTGAAGAGGT 1019
```

RESULT 11
US-10-437-963-15200
; Sequence 15200, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Query Match 7.3%; Score 181; DB 7; Length 1118;
Best Local Similarity 73.2%; Pred. No. 3.7e-38;
Matches 232; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
QY 642 ACATAGGCGAGGTAGAGCGGTCCATCGCTTACAGGAGGTATTTGCTGTGATTT 701
DB 61 ACGAAGGTGAGAGCTGCATCAGTTGTACCATCTCTACACAACGATTTCTCTCGGTGATG 120
QY 702 GTGAGAGAGATTTCTGCAAAATGTGAGAGAGGCTTTGTGAGAAAGTTGCAGCCTTATGCA 761
DB 121 GTTGACAGAGTTTTTTGTAATGTCTGAGAGATTAGTGGAGAGCTTGAGACATCTGCT 180
QY 762 GAAGACGGAAGTGTGTGAATATGGAAGCGAAAGTTCTCTCAGATGACATTTGATGCTATT 821
DB 181 TTAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATGACTTTTAGATGTTGATT 240
QY 822 GGGTTGTCTCTTTTAACTTACAAATTTGATTTCTTTGACTACTGATGATCTCTGTCATTGAA 881
DB 241 GGTGTGCTCTTGTTCATTTACATTTTGATTTCCCTCAGATAGAGCCCTGTTATTGAT 300
QY 882 GCTGTTTACACTGCTCTTAAAGAGCTGAGCTTTCGTTCTACTGATCTTCTGCCATATTGG 941
DB 301 GCTGTTTACACTGCACTCAAGGAGCAGAACTTCGTTCTACAGATCTTTTACCATCTGS 360
QY 942 AAGGCAAGTTTCCCTGTG 958
DB 361 AAGATTGATTTGCTGTG 377
```

RESULT 12
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li

APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-10-751-235-27

Query Match 6.9%; Score 169.6; DB 9; Length 624;
Best Local Similarity 69.6%; Pred. No. 3.4e-35;
Matches 245; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

```
QY 608 TTCTCTATTTGGGCTTAGCTGAAGTCTGATTTTTTACATAGGCGAGCGGTAGAGCGGTGGT 667
DB 117 TTCTGGTTTGGCAATTGCTGAAGGCTCTCTTTGGAC---TGCAAGCGCGAGGCTGTAGT 173
QY 668 TCCATCGCTTACAGAGGATTTTCTGTGTTTGTGAGAGATTTCTTGCANATGTGC 727
```

Db 174 TCCATCACTTCAAGAGAGTACTTATCAGTAATAGTGTGATGCTGATTTTGCAGATGCTC 233
QY 728 AGAGAGGCTTGTGAGAGTTGACGCCCTTATGACAGAGGAGGAGTGTGTAATATGGA 787
Db 234 CGAGAGGCTTGTGAGAGAGTAAAGATCATACGACGAGTGCACAGTCTGTGTAACATGGA 293
QY 788 AGCGAAGTCTCTCAGATGACACTTGTATGATGCTATGCTGCTCTCTTTTAACTACAATT 847
Db 294 GCAACAGTTTTCGCGAGTTAAACCTTGATGTTATCGGCTAGCCGTAATTTAACTACAATT 353
QY 848 CGATCTTTGACTACTGATAGTCCCTGTCATTTGAAGCTGTTTACACTGCTCTTAAAGAGC 907
Db 354 TGACTCACTTACGCGGATAGTCCCTGTAATGTAATCTGTTTATACCGCACTAAAGAGC 413
QY 908 TGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGCAAGTTTCCCTGCTGT 959
Db 414 TGAAGCCGTTCAACTGATCTTTTGCCATATTTGGAAGATAGTGGCTTATGT 465
RESULT 13
US-10-751-235-25
; Sequence 25, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 839
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-10-751-235-25
Query Match 6.5%; Score 161; DB 9; Length 839;
Best Local Similarity 61.7%; Pred. No. 9.2e-33;
Matches 329; Conservative 0; Mismatches 115; Indels 89; Gaps 1;
QY 717 TCAAAATGTGCAGAGAGGCTTGTGAGAAAGTTGCAGCCTTATGACAGAGCGGAAGTGT 776
Db 1 TGCAGATGTGCTGAGAGAAATGGTGGAGAACTTTTACCTGATGCAATTTCTGCTCTGCA 60
QY 777 GTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGATGCTCAATTGGGTTGCTCTTTT 836
Db 61 GTGAATATGGAAGCGAAGTTTCTCAACTAACAATGATGTTATTTGGCCTTGCACTCTTC 120
QY 837 AACTACAAATTTGATTTCTTGTACTGATGATGCTCTGCTCAATTTGAAGCTGTTTACACTGCT 896
Db 121 AATTACAAATTTGATTTCCCTTACTACTGACAGTCCAGTATTTGATGTCAGTTTACACTGCA 180
QY 897 CTTAAAGAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCATATTTGCCATATTTTCCCTG 956
Db 181 CTTAAAGAGAGCAAACTCCGTTCAACTGATTTGTTGCCATATTTGG----- 225
QY 957 TGTTTTTCTGTGGTTTGTGATGTGTGGAACAATTTGATTTCTTGTTAATTGAGAGGTT 1016
Db 226 ----- 225
QY 1017 TTGGTTGTTTTTTTTCAGATCGATGTCATTTGTGTAAGATAGTCCCGACAGAGGTGAAAGCTG 1076
Db 226 -----CAGATCAAAGCTTTATGTAAGTTTCATCCACGACAAATPAAAGGCTG 271
QY 1077 AAAAGGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTATTTGCTAAGTGTAAAGAAA 1136
Db 272 AGAATGCAAGTGTCAATTAATCAGACAAACAGTTGAAGAACTTATTTGCGAAGTGCAGAGAGA 331
QY 1137 TTGTGCAAGAGAGCGGAAGAAATCAATGATGAGGATGTAAATGATGCTGACCCAA 1196

Db 332 TTGTAGAAACTGAGGCTGAGAGGATTAATGAAGATAGTACGTGAATGATAGATCCAA 391
QY 1197 GTATCTCGCTTTCTTCTGCTCAAGCAGAGAGAGGTTTAAACTTTTTCCTT 1249
Db 392 GCATCTCTCGAATTTTGTCTGTAGCCGTGAGGAGGTTTCAAGTTTACAACCTT 444
RESULT 14
US-10-751-235-26
; Sequence 26, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Zea mays
US-10-751-235-26
Query Match 5.9%; Score 144.6; DB 9; Length 531;
Best Local Similarity 62.1%; Pred. No. 2.2e-28;
Matches 300; Conservative 0; Mismatches 94; Indels 89; Gaps 1;
QY 751 AGCCTTATGCAGAACGAGAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATGACAC 810
Db 5 AGCCATATGCTTTGAGTGGGAAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTGACAT 64
QY 811 TTGATGTCAATTTGGGTTGTCTCTTTTAACTACAATTTTCGATTTCTTGTACTACTGATGTC 870
Db 65 TGGATGTGATTTGGTTTATTCATTTGTTCACTACAAATTTTGAATTTCCCTCACAACAGATGTC 124
QY 871 CTGTCAATTTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTC 930
Db 125 CTGTCAATTTGATGCTGTTTATATCTGCACCTCAAAGAGCAGAGCTTCGTTCTACAGATCTTT 184
QY 931 TGCCATATTTGGAAGCGAAGTTTCTGCTGTGTTTTTCTGTGTTTGTGTTGTTGTGGAACA 990
Db 185 TGCCATATCTGGAAGGTTGGTTTC----- 207
QY 991 ATTTGGATTTCTGTTAATTTGAGAGGGTTTGGTTGTTTTTTTTCAGATCGATGCTGTTGTA 1050
Db 208 -----TTGTGCAA 215
QY 1051 GATAGTCCGACAGAGGTGAAAGCTGAAAGGCTGTAACTTTGATTAAGGAAACTGTTGA 1110
Db 216 GATAATCCCAAGACAGATATAAGCAGAGAAATGCGGTAGCAATATATAAGGAACACTGTTGA 275
QY 1111 AGACCTTATTTGCTAAGTGTAAAGAAATTTCTCCAAAGAGAGCGGAAGATCAATCATGA 1170
Db 276 AGAGCTGATTTATTAAGTGTAAAGAAATATGTTGGAAGCTGAAATGAATGAACAGATTGAGG 335
QY 1171 GGAGTATGTAAATGATGCTGACCCAGTATCTTGCCTTTCTTGTCTTCTTGTGCAAGCAGAGA 1230
Db 336 GGAATATGTAACGAAGGGATCTTAGCAATTTCTACGTTCTTCTTGTGCTAGCCGAGATGA 395
QY 1231 GGT 1233
Db 396 GGT 398
RESULT 15
US-10-437-963-15203
; Sequence 15203, Application US/10437963

/ Publication No. US20040123343A1

/ GENERAL INFORMATION: 63.8%; Score 128.8; DB 7; Length 1596;

/ APPLICANT: La Rosa, Thomas J.

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ APPLICANT: Wu, Wei

/ APPLICANT: Boukharov, Andrey A.

/ APPLICANT: Barbazuk, Brad

/ APPLICANT: Li, Ping

/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53221)B

/ CURRENT APPLICATION NUMBER: US/10/437,963

/ NUMBER OF SEQ ID NOS: 204966

/ SEQ ID NO 15203

/ LENGTH: 1596

/ TYPE: DNA

/ ORGANISM: Oryza sativa

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MRT4530_21069C.1

US-10-437-963-15203

Query Match 5.2%; Score 128.8; DB 7; Length 1596;

Best Local Similarity 63.8%; Pred. No. 8.6e-24;

Matches 231; Conservative 0; Mismatches 122; Indels 9; Gaps 2;

```
QY 158 CCCATCATGGTCCGATTGGCTCACAACACTCACTCGTACCCCTTTC-----CT 211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 CAGGAAAAAAGCAGAGTCAGGTATACCAATCCGAACGCGAAGCTCGACGATGCGCTG 271
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 TCGCGGGGGCGGACGACTCGGGGATCCCGTCCGCTCCGCAAGCTCGACGACGTGCGGG 292
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 ATCTCTCGGAGGTGCTCTCTTCTTACCTCTCTACAAATGGATGAATGAGTACGACCCA 331
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 ACCTCTCGGGCGCGCTCTTCTCTCCCTCTCTCAAGTGTTCGCGGAGGAAGGCCCG 352
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 TTACCGTCTCGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 TCTACCGCTCGCGGGGGGGCGCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 AACATGTTTGT---AGGAATTATCCAAAGTACGCTAAAGGCTTAGTCTGTAAGTCTCTG 448
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 GGCAGTGTCTGCTGGGTACGGTTCGAGGTACGAGAGGGGCTCGTGGCGAGGTTTCCG 472
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 449 AATTCTATTGGTTTCGGGTTTCGCTATCGCTGAAGGACCTCTTTGGACAGTAATTTCAT 508
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 AGTTCCTCTTCGGCTCCGGGTTCCGCATCGCGAGGGGCTCTCTGGACGTTTCTCGTAT 532
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 509 CT 510
Db 533 GT 534
```

Search completed: January 13, 2006, 08:04:03

Job time : 1962 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 19:33:24 ; Search time 1445 Seconds
(without alignments)
11378.406 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagttctactcttttc.....tgaagtgagccaaaggttaa 2467

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21.1*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	100.0	2467	14 AEB16909	Aeb16909 Thale cre
2	2467	100.0	4170	14 AEB16910	Aeb16910 Thale cre
3	2465.4	99.9	2467	14 AEB16911	Aeb16911 Thale cre
4	803	32.5	3377	13 ADT15447	Adt15447 Plant cDN
5	195.8	7.9	1804	13 ADX61765	Adx61765 Plant ful
6	187.2	7.6	1086	14 AEB16926	Aeb16926 Wheat cyt
7	187.2	7.6	1638	14 AEB16925	Aeb16925 Wheat cyt
8	183	7.4	1686	14 AEB16924	Aeb16924 Rice cyto
9	169.6	6.9	624	14 AEB16929	Aeb16929 Sunflower
10	161	6.5	839	14 AEB16927	Aeb16927 Tomato cy
11	144.6	5.9	531	14 AEB16928	Aeb16928 Maize cyt
12	91	3.7	1788	14 AEB16943	Aeb16943 Thale cre
13	91	3.7	2057	14 AEB16942	Aeb16942 Thale cre
14	90.8	3.7	1278	14 AEB16947	Aeb16947 Soybean c
15	87.4	3.5	1078	13 ADO82389	Ado82389 Plant ful
16	84.6	3.4	1737	14 AEB16949	Aeb16949 Tomato cy
17	83.4	3.4	1899	14 AEB16945	Aeb16945 Rice cyto
18	76	3.1	2239	13 ADO81358	Ado81358 Plant ful
19	73.2	3.0	1731	14 AEB16957	Aeb16957 Soybean c

20	73.2	3.0	1831	2 AAX60780	Aax60780 Soybean c
21	72.8	3.0	1031	14 AEB16948	Aeb16948 Wheat cyt
22	72.8	3.0	1527	14 AEB16946	Aeb16946 Barley cy
23	72.8	3.0	1926	14 AEB16956	Aeb16956 Pea CYP97
24	72	2.9	1734	14 AEB16955	Aeb16955 Thale cre
25	68.8	2.8	1101	14 AEB16950	Aeb16950 Green alg
26	65.2	2.6	5071	14 AEB16944	Aeb16944 Rice cyto
27	59.4	2.4	579	10 ABX56831	Abx56831 Arabidops
28	57.8	2.3	404	3 AAA68063	Aaa68063 Pinus rad
29	57.8	2.3	404	10 ADD41813	Add41813 Cinnammat
30	56.6	2.3	1980	14 AEB16959	Aeb16959 Diatom cy
31	51.8	2.1	1299	13 ADS57092	Ads57092 Bacterial
32	51.2	2.1	1321	13 ADT44250	Adt44250 Bacterial
33	51	2.1	2000	8 ADA71938	Ada71938 Rice gene
34	50	2.0	297	12 ADN37071	Adn37071 Tobacco c
35	50	2.0	313	12 ADO85319	Ado85319 Nicotiana
36	50	2.0	313	12 ADO85355	Ado85355 Nicotiana
37	50	2.0	313	12 ADO85355	Ado85355 Nicotiana
38	50	2.0	1595	12 ADN10644	Adn10644 Nicotiana
39	50	2.0	1595	14 ADZ65332	Adz65332 Tobacco f
40	50	2.0	1595	14 ADZ65879	Adz65879 Tobacco c
41	49.2	2.0	2000	11 ACL35887	Acl35887 Rice stre
42	48.6	2.0	2000	11 ACL37108	Acl37108 Rice stre
43	47.8	1.9	259	12 ADN37087	Adn37087 Tobacco c
44	47.8	1.9	259	12 ADO85335	Ado85335 Nicotiana
45	47.6	1.9	2000	8 ADA71938	Ada71938 Rice gene

ALIGNMENTS

RESULT 1

AEB16909
ID AEB16909 standard; cDNA; 2467 BP.

AC AEB16909;

DT 08-SEP-2005 (first entry)

XX Thale cress LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 5.

XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
XX cytochrome P450 97C1; gene; ss; chromosome 3.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers
XX CDS 1..2467

FT /product= "Thale cress LUT1 protein"
FT /transl_except= (pos:496..650, aa:Thr-Ala)
FT /transl_except= (pos:942..1037, aa:Lys-Ile)
FT /transl_except= (pos:1242..1326, aa:Val-Ser)
FT /transl_except= (pos:1420..1523, aa:Lys-Asn)
FT /transl_except= (pos:1662..1841, aa:Val-Leu)
FT /transl_except= (pos:1839..2005, aa:Val-Trp)
FT /transl_except= (pos:2081..2183, aa:Lys-Phe)
FT /transl_except= (pos:2340..2440, aa:Asn-Gly)

US2005150002-A1.

07-JUL-2005.

02-JAN-2004; 2004US-00751235.

02-JAN-2004; 2004US-00751235.

(DELL/) DELLAPENNA D.

(TIAN/) TIAN L.

(KIMJ/) KIM J.

Dellapenna D, Tian L, Kim J;

XX

= instant
application


```
QY 1801 TTATGATTTTCAGTCTATATACATCCATCGTTCTTCGAGGTACAGTTCTCTCTCTCTCTC 1860
Db 1801 TTATGATTTTCAGTCTATATAACATCCATCGTTCTTCGAGGTACAGTTCTCTCTCTCTC 1860
QY 1861 GTCCATAGTATATACATAGGGAGCCCTAAATCCCTCTCTTCAATGATCTTTGTGTGGTTCGG 1920
Db 1861 GTCCATAGTATATACATAGGGAGCCCTAAATCCCTCTCTTCAATGATCTTTGTGTGGTTCGG 1920
QY 1921 ATATCTAACCGAGTGGACATCTCTAGTATTAATCATTCAGGCCACATTTCTTATGTGTTT 1980
Db 1921 ATATCTAACCGAGTGGACATCTCTAGTATTAATCATTCAGGCCACATTTCTTATGTGTTT 1980
QY 1981 GTTGTGTTGTTATTTCCAAAGGTATGGAAAAAGCTGAGGAATTTCTGCCCTGAACGATTCGA 2040
Db 1981 GTTGTGTTGTTATTTCCAAAGGTATGGAAAAAGCTGAGGAATTTCTGCCCTGAACGATTCGA 2040
QY 2041 CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC 2100
Db 2041 CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC 2100
QY 2101 ATCTTTTGACACAAACTACTGAATCAAGATTAGTGTGTTTGTGATTTAGGGAATTTAAAGAT 2160
Db 2101 ATCTTTTGACACAAACTACTGAATCAAGATTAGTGTGTTTGTGATTTAGGGAATTTAAAGAT 2160
QY 2161 GATTTTCTTTTTCACCAATTCATCCCAATTCAGTGAGGGGCTAGAAATGCTGAGCG 2220
Db 2161 GATTTTCTTTTTCACCAATTCATCCCAATTCAGTGAGGGGCTAGAAATGCTGAGCG 2220
QY 2221 ATCAGTTTGCAATGATGAGGCAATTTGTGGCACTCGCGGTGTTCTTACGCGGTTAAAG 2280
Db 2221 ATCAGTTTGCAATGATGAGGCAATTTGTGGCACTCGCGGTGTTCTTACGCGGTTAAAG 2280
QY 2281 TTGAGCTGGTTCCTGATCAGACCAATAGCATGACCAACAGGACCAACATACACACCA 2340
Db 2281 TTGAGCTGGTTCCTGATCAGACCAATAGCATGACCAACAGGACCAACATACACACCA 2340
QY 2341 ATGATATGCCAATGTTCTACACTCGAGATTAATGAGAGTGTCTGTTTGTGTTAGATG 2400
Db 2341 ATGATATGCCAATGTTCTACACTCGAGATTAATGAGAGTGTCTGTTTGTGTTAGATG 2400
QY 2401 ATTCCAATTTCTTAATGCTGATATTTTCAATTTTCAGGGATGTTATGAAAGGTGAGCA 2460
Db 2401 ATTCCAATTTCTTAATGCTGATATTTTCAATTTTCAGGGATGTTATGAAAGGTGAGCA 2460
QY 2461 AAGGTAA 2467
Db 2461 AAGGTAA 2467
```

RESULT 2

```
AEBl6910
ID AEB16910 standard; DNA; 4170 BP.
AC AEB16910;
XX
XX
DT 08-SEP-2005 (first entry)
DE Thale cress LUT1 (cytochrome P450 97C1) DNA, SEQ ID NO: 6.
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
KW cytochrome P450 97C1; gene; db; chromosome 3.
XX
XX OS Arabidopsis thaliana.
XX PN US2005150002-A1.
XX
XX PD 07-JUL-2005.
XX
XX PF 02-JAN-2004; 2004US-00751235.
XX
XX PR 02-JAN-2004; 2004US-00751235.
XX
XX PA (DELL/) DELLAPENNA D.
```

```
PA (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
PI Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 6; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is thale cress LUT1 (cytochrome P450
CC monooxygenase (CYP97C1); At3g5130 gene) genomic DNA. Note: The current
CC sequence is that of thale cress LUT1 DNA which is located on chromosome
CC 3.
XX
SQ Sequence 4170 BP; 1222 A; 862 C; 812 G; 1274 T; 0 U; 0 Other;
Query Match 100.0%; Score 2467; DB 14; Length 4170;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGTCTTCACTCTTTTCTCCATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 60
Db 1010 ATGAGTCTTCACTCTTTTCTCCATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 1069
QY 61 CCTACGCTCTTTTATACCAAAACCAAAATTCATCGAAATCCCAATCATGAGTCCAGTCCCGAT 120
Db 1070 CCTACGCTCTTTTATACCAAAACCAAAATTCATCGAAATCCCAATCATGAGTCCAGTCCCGAT 1129
QY 121 AAACCCAAACCAAACTCGAGACCAATTCATCGAAATCCCAATCATGAGTCCAGTCCCGAT 180
Db 1130 AAACCCAAACCAAACTCGAGACCAATTCATCGAAATCCCAATCATGAGTCCAGTCCCGAT 1189
QY 181 TGGCTCAACAACACTCACTCGTACCTTTCTCAGAGAAAACGACGAGTACGATATACCA 240
Db 1190 TGGCTCAACAACACTCACTCGTACCTTTCTCAGAGAAAACGACGAGTACGATATACCA 1249
QY 241 ATCGGACGCGAGAGCTCGAGATGCTCGATCTCTCGGAGGTCTCTTCTTACCT 300
Db 1250 ATCGGACGCGAGAGCTCGAGATGCTCGATCTCTCGGAGGTCTCTTCTTACCT 1309
QY 301 CTCTACAAATGGATGAATGAGTACGAGACCCATTTTACCGTCTCGTCTGCTGCTGCTGCTGCTGCT 360
Db 1310 CTCTACAAATGGATGAATGAGTACGAGACCCATTTTACCGTCTCGTCTGCTGCTGCTGCTGCTGCT 1369
QY 361 TTGCTAATTTGAGCGACCCAGCGATAGCTAAACATATGTTTGGAGAAATATCCAAAGTAC 420
Db 1370 TTGCTAATTTGAGCGACCCAGCGATAGCTAAACATATGTTTGGAGAAATATCCAAAGTAC 1429
QY 421 GCTAAGGCTTAGTGGTGAAGTCTCTGAATTTCTATTTGGTTCGGGTTTCGCTATCGCT 480
Db 1430 GCTAAGGCTTAGTGGTGAAGTCTCTGAATTTCTATTTGGTTCGGGTTTCGCTATCGCT 1489
QY 481 GAAGGACCTCTTTGGACAGTAAATTTCAATCTCTCTCTATCTCAATTTTGAAGTTTGTGAA 540
Db 1490 GAAGGACCTCTTTGGACAGTAAATTTCAATCTCTCTCTATCTCAATTTTGAAGTTTGTGAA 1549
QY 541 TTGTGGAAGTAAATGTGACGTCTGTGTATGATAGTAAGTAACTCTAATTTAGGTTTAGAT 600
Db 1550 TTGTGGAAGTAAATGTGACGTCTGTGTATGATAGTAAGTAACTCTAATTTAGGTTTAGAT 1609
QY 601 TCCATCTTCTTATTTGGGCTTAGCTGAAGTCTGATTTTTCATAGGAGGCGGTAGAG 660
Db 601 TCCATCTTCTTATTTGGGCTTAGCTGAAGTCTGATTTTTCATAGGAGGCGGTAGAG 660
```

Db 1610 TCCAACTCTCTCTATTGGGCTTAGCTGAAGTCTGATTTTTTACATAGCGAGGGGTAGAG 1669
Qy 661 CGGTGGTTCCATCGCTTCAAGGAGGTATTTGTCTGTGATTTGGAGAGATATTTCTGCA 720
Db 1670 CGGTGGTTCCATCGCTTCAAGGAGGTATTTGTCTGTGATTTGGAGAGATATTTCTGCA 1729
Qy 721 AATGTGAGAGAGGGTTGTGAGAAAGTTGCAGCCTTATGCGAGACCGAAGTGTCTGGA 780
Db 1730 AATGTGAGAGAGGGTTGTGAGAAAGTTGCAGCCTTATGCGAGACCGAAGTGTCTGGA 1789
Qy 781 ATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCAATTTGGGTGTCTCTTTTAACT 840
Db 1790 ATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCAATTTGGGTGTCTCTTTTAACT 1849
Qy 841 ACAATTTCCGATTTCTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGCTCTTA 900
Db 1850 ACAATTTCCGATTTCTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGCTCTTA 1909
Qy 901 AAGAAGCTGAGCTTCGTTCTACTGATCTCTGCCATATTTGGAAGCAAGTTTCCGTGTT 960
Db 1910 AAGAAGCTGAGCTTCGTTCTACTGATCTCTGCCATATTTGGAAGCAAGTTTCCGTGTT 1969
Qy 961 TTTTCTGTGTTTGTGATTTGTGTGGAACAAATTTGGAATCTCTGTTAAATGAGAGGGTTTGG 1020
Db 1970 TTTTCTGTGTTTGTGATTTGTGTGGAACAAATTTGGAATCTCTGTTAAATGAGAGGGTTTGG 2029
Qy 1021 TTGTTTTTTTTCAGATCGATGCAATTTGTGAAGATAGTCCCGAGACAGGTGAAGCTGGAATA 1080
Db 2030 TTGTTTTTTTTCAGATCGATGCAATTTGTGAAGATAGTCCCGAGACAGGTGAAGCTGGAATA 2089
Qy 1081 GCGTGTAACTTTGATAAGGAACTGTTGAAGACCTTATTGTAAGTGTGAAGAAATTTGT 1140
Db 2090 GCGTGTAACTTTGATAAGGAACTGTTGAAGACCTTATTGTAAGTGTGAAGAAATTTGT 2149
Qy 1141 CGAAGAGAGGCGAAGAACTCAATGATGAGGAGTATGTAATGATGCTGACCCAAAGTAT 1200
Db 2150 CGAAGAGAGGCGAAGAACTCAATGATGAGGAGTATGTAATGATGCTGACCCAAAGTAT 2209
Qy 1201 CTGCGGTTTCTTGTGCAAGCAGAGAGAGGTTTAAACTTTTTTCTTTAAGTTTATAAG 1260
Db 2210 CTGCGGTTTCTTGTGCAAGCAGAGAGAGGTTTAAACTTTTTTCTTTAAGTTTATAAG 2269
Qy 1261 CAAATTTGGCCCTTCAATTCGATTAATCGAAGCTGATGTTGCATTTGAGGAGGTTTTCAG 1320
Db 2270 CAAATTTGGCCCTTCAATTCGATTAATCGAAGCTGATGTTGCATTTGAGGAGGTTTTCAG 2329
Qy 1321 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCATGAAACC 1380
Db 2330 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCATGAAACC 2389
Qy 1381 ACTGGATCTGTCTCACTTTGGACACTTTATCTCTTAAGTAAAGTACCTTAATGTATCTTC 1440
Db 2390 ACTGGATCTGTCTCACTTTGGACACTTTATCTCTTAAGTAAAGTACCTTAATGTATCTTC 2449
Qy 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTGTGTTCTCATTTTACCTCTT 1500
Db 2450 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTGTGTTCTCATTTTACCTCTT 2509
Qy 1501 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAAGCAAGAAAGATGAGACA 1560
Db 2510 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAAGCAAGAAAGATGAGACA 2569
Qy 1561 GAGTGTTAGAAGGAAGAAACCCGGCTTTCAGAGATATAAGGAGTTGAAGTACATCACTC 1620
Db 2570 GAGTGTTAGAAGGAAGAAACCCGGCTTTCAGAGATATAAGGAGTTGAAGTACATCACTC 2629
Qy 1621 GTTGTATAAACGAGTCAATGCGTCTCTATCCTCATCTCTGTGAAGCAATCAAGCTCATC 1680
Db 2630 GTTGTATAAACGAGTCAATGCGTCTCTATCCTCATCTCTGTGAAGCAATCAAGCTCATC 2689
Qy 1681 TCTCTAATTTATTCATGAACCTAAATTTTCTGATTTGTTTCTCGTAGGTTCTTGATAA 1740
Db 2690 TCTCTAATTTATTCATGAACCTAAATTTTCTGATTTGTTTCTCGTAGGTTCTTGATAA 2749

Qy 1741 GAAGAGCTCAAGTTCCCTGACATTTCTTCTGGGAACCTATAAGGTCAATACCGGACAAGACA 1800
Db 2750 GAAGAGCTCAAGTTCCCTGACATTTCTTCTGGGAACCTATAAGGTCAATACCGGACAAGACA 2809
Qy 1801 TTATGATTTCAAGTCTATAAACAATCCATCGTTCTTCCGAGGTACAGTTCTCTTCTCTC 1860
Db 2810 TTATGATTTCAAGTCTATAAACAATCCATCGTTCTTCCGAGGTACAGTTCTCTTCTCTC 2869
Qy 1861 GTCCATAGTATAACATAGGGAGGCTTAATCTCTTCTCAATGATCTTTGTGTTGTTGCG 1920
Db 2870 GTCCATAGTATAACATAGGGAGGCTTAATCTCTTCTCAATGATCTTTGTGTTGTTGCG 2929
Qy 1921 ATATCTAACCGGAGTGAACATTTCTAGTATTTACATTTCAATGCCACATTTCTTATGTGTTT 1980
Db 2930 ATATCTAACCGGAGTGAACATTTCTAGTATTTACATTTCAATGCCACATTTCTTATGTGTTT 2989
Qy 1981 GTTGTGTTGTTATTTCCAAAGGTATGGAAGGCTGAGGAAATTTCTGCTGAACGATTCGA 2040
Db 2990 GTTGTGTTGTTATTTCCAAAGGTATGGAAGGCTGAGGAAATTTCTGCTGAACGATTCGA 3049
Qy 2041 CATAGATGGCGCAATCCCTTAACGAAACAACAACATGATTTCAAGTAAACTCAGTAGAACAC 2100
Db 3050 CATAGATGGCGCAATCCCTTAACGAAACAACAACATGATTTCAAGTAAACTCAGTAGAACAC 3109
Qy 2101 ATCTTTTTCACACAAACTACTGAACTCAAGATTTAGTGGTTTGTGATTTAGGGAAATTTAAAGAT 2160
Db 3110 ATCTTTTTCACACAAACTACTGAACTCAAGATTTAGTGGTTTGTGATTTAGGGAAATTTAAAGAT 3169
Qy 2161 GATTTCTTTTTCACACAGATTTCAATCCATTCAGTGGAGGGCTAGAAAATGTGTAGGG 2220
Db 3170 GATTTCTTTTTCACACAGATTTCAATCCATTCAGTGGAGGGCTAGAAAATGTGTAGGG 3229
Qy 2221 ATCAAGTTGCAATGATGAGGCAATTTGGCACTCGCGGTGTTTCTTCAAGCGGTTAAACG 2280
Db 3230 ATCAAGTTGCAATGATGAGGCAATTTGGCACTCGCGGTGTTTCTTCAAGCGGTTAAACG 3289
Qy 2281 TTGAGCTGTTCTCATCAGACCAATTTAGCATCAGGAGCAACCAACCAACCAACCAACCA 2340
Db 3290 TTGAGCTGTTCTCATCAGACCAATTTAGCATCAGGAGCAACCAACCAACCAACCAACCA 3349
Qy 2341 ATGTATGCCAAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCTGTTTGTGTTGAATG 2400
Db 3350 ATGTATGCCAAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCTGTTTGTGTTGAATG 3409
Qy 2401 ATTCCAATTTCTTAATGCTGATATTTCAATTTCCAGGATTTGATATGAAGTGAAGCCA 2460
Db 3410 ATTCCAATTTCTTAATGCTGATATTTTCAATTTCCAGGATTTGATATGAAGTGAAGCCA 3469
Qy 2461 AAGGTAA 2467
Db 3470 AAGGTAA 3476

RESULT 3

ABE16911
ID ABE16911 standard; cDNA; 2467 BP.

XX ABE16911;

XX AC ABE16911;

XX DT 08-SEP-2005 (first entry)

XX DE Thale cress mutant LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 7.

XX KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1; cytochrome P450 97C1; ss; mutant.

XX OS Arabidopsis thaliana.

XX OS Synthetic.

XX XX US2005150002-A1.

XX PN 07-JUL-2005.

QY 61 CCTACGCTCTTTTATCACCAAAACCCAAATTCATTCATCAGATCCTCCATTGAG 120
DB 139 CCTACGCGTCTTTTATCACAAACCCAAATTCATTCATCAGATCCTCCATTGAG 198
QY 121 AAACCCAAACCCAAACCTCGAGACCAATTCATCGAAATCCCAATCATGGGTGATGCCGAT 180
DB 199 AAACCCAAACCCAAACCTCGAGACCAATTCATCGAAATCCCAATCATGGGTGATGCCGAT 258
QY 181 TGGCTCAACAACCTCATCGTACCTTTCTCGAGGAAACCAACGACGAGTCAGGTATACCA 240
DB 259 TGGCTCAACAACCTCATCGTACCTTTCTCGAGGAAACCAACGACGAGTCAGGTATACCA 318
QY 241 ATCGGGAACCGAAGCTCGACGATGCTGATCTCTCGAGGTGCTCTCTCTTACCT 300
DB 319 ATCGGGAACCGAAGCTCGACGATGCTGATCTCTCGAGGTGCTCTCTCTTACCT 378
QY 301 CTCTACAAATGGATGAATGAGTACGACCAATTTACCGTCTCGCTGCTGCTCGTAAAT 360
DB 379 CTCTACAAATGGATGAATGAGTACGACCAATTTACCGTCTCGCTGCTGCTCGTAAAT 438
QY 361 TTCTGTAATTTGAGGACCCAGGATGCTAAACATGTTTTGAGGAATTTATCCAAAGTAC 420
DB 439 TTCTGTAATTTGAGGACCCAGGATGCTAAACATGTTTTGAGGAATTTATCCAAAGTAC 498
QY 421 GCTTAAGGCTTAGTCTGCTGAGCTCTGAAATTTCTATTTGTTTGGGTTTCGCTATCGCT 480
DB 499 GCTTAAGGCTTAGTCTGCTGAGCTCTGAAATTTCTATTTGTTTGGGTTTCGCTATCGCT 558
QY 481 GAAGGACCTCTTTTGGACAGTAATTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 559 GAAGGACCTCTTTGGACA 576
QY 541 TTGTGGAGTAATGTTGATGCTCTGTTGATGATAAGTAACTCTAATTTTGGGTTTAGAT 600
DB 577 ----- 576
QY 601 TCCAACTCTCTATTTGGGCTTAGCTGAAGTCTGATTTTATACATAGGCGAGCGTAGAG 660
DB 577 -----GCGAGGCGTAGAG 589
QY 661 CGGTGGTTCCATCGCTTTCACAGAGGATTTCTGCTGATTTGTTGAGAGAGTATTTCTGCA 720
DB 590 CGGTGGTTCCATCGCTTTCACAGAGGATTTCTGCTGATTTGTTGAGAGAGTATTTCTGCA 649
QY 721 AATGTGCAGAGAGGCTTTGTGAGAAGTTGACGCTTATGACAGAGACGGAAGTCTGTGA 780
DB 650 AATGTGCAGAGAGGCTTTGTGAGAAGTTGACGCTTATGACAGAGACGGAAGTCTGTGA 709
QY 781 ATATGGAGGAGGATTTCTCTGATGACACTTTGATGCTATTTGGGTTGCTCTTTTAACT 840
DB 710 ATATGGAGGAGGATTTCTCTGATGACACTTTGATGCTATTTGGGTTGCTCTTTTAACT 769
QY 841 ACAATTTTGGATTTTGTACTGATGATGCTGCTGATGCTGCTGATGCTGCTGCTGCTGCT 900
DB 770 ACAATTTTGGATTTTGTACTGATGATGCTGCTGATGCTGCTGATGCTGCTGCTGCTGCT 829
QY 901 AAGAAGCTGAGCTTCTGTTCTACTGATCTTTCTGCTATTTGGAAGGCAAGTTTCTGTGTT 960
DB 830 AAGAAGCTGAGCTTCTGTTCTACTGATCTTTCTGCTATTTGGA 871
QY 961 TTTTCTGTGTTTGTGATTTGTGTGGAACAATTTGATTTCTGTTAATTTGAGAGGTTTGG 1020
DB 872 ----- 871
QY 1021 TTGTTTTTTTTCAGATGCTATTTGTGAAGATAGTCCCGACAGAGTGAAGCTGAATA 1080
DB 872 -----AGATGCTATTTGTGAAGATAGTCCCGACAGAGTGAAGCTGAATA 920
QY 1081 GGTCTGTAATTTGATAAGGGAACCTGTTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTTGT 1140
DB 921 GGTCTGTAATTTGATAAGGGAACCTGTTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTTGT 980
QY 1141 CGAAAGAGAGGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTAT 1200

DB 981 CGAAAGAGAGGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTAT 1040
QY 1201 CCTGGGTTTCTTTGCTTTGCAAGCAGAGAGGTTTAAACTTTTTTCTTTAAGTTTATAAG 1260
DB 1041 CCTGGGTTTCTTTGCTTTGCAAGCAGAGAG- 1069
QY 1261 CAAATTTGGCCTTTTCATTTATCGCAATAATCGAAGCTGATGTTGCAATTTGAGGGTTTTCAG 1320
DB 1070 -----AG 1071
QY 1321 GTATCAAGTGTGCTGAGTTACGGGATGATCTTCTCAATGCTCTAGCGGTCATGAAC 1380
DB 1072 GTATCAAGTGTGCTGAGTTACGGGATGATCTTCTCAATGCTCTAGCGGTCATGAAC 1131
QY 1381 ACTGGATCTGCTCTCACTTGGACACTTTATCTCTTAAGTAAGGTACTTTAATGTATCTTC 1440
DB 1132 ACTGGATCTGCTCTCACTTGGACACTTTATCTCTTAAGTA 1171
QY 1441 TACTTTGCTATGCTAGAGAAATTTA CTGGATGGGAGCTTCTCTGTTCTCATTTACCTCTT 1500
DB 1172 ----- 1171
QY 1501 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAACCAAGAGAGTAGACA 1560
DB 1172 -----AGAACTCATCTGCAATTAAGGAAACCAAGAGAGTAGACA 1213
QY 1561 GAGTGTTAGAAGAAAGAAACCCGGCTTTTCGAGGATATAAGGAGTTGAAGTACATCACCTC 1620
DB 1214 GAGTGTTAGAAGAAAGAAACCCGGCTTTTCGAGGATATAAGGAGTTGAAGTACATCACCTC 1273
QY 1621 GTTGTATAAAACGAGTCAATGCTCTATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
DB 1274 GTTGTATAAAACGAGTCAATGCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1314
QY 1681 TCTCTAATTTATTCAGAACTAAATTTTCTGATGATTTGTTTCTCTGTTAGTCTTTGATAA 1740
DB 1315 -----GTTCTTGATAA 1324
QY 1741 GAAGAGCTCAAGTTCTCTGACATTTCTTCTCTGGAACTATAAGGTCAATACCGACAGACA 1800
DB 1325 GAAGAGCTCAAGTTCTCTGACATTTCTTCTCTGGAACTATAAGGTCAATACCGACAGACA 1384
QY 1801 TTATGATTTCACTCTATTAACATCCATCCATCGTTCTTCTCCGAGGTA 1841
DB 1385 TTATGATTTCACTCTATTAACATCCATCCATCGTTCTTCTCCGAGGTA 1425

RESULT 5

ADX61765

ID ADX61765 standard; cDNA; 1804 BP.

XX

AC ADX61765;

XX

DT 21-APR-2005 (first entry)

XX

DE Plant full length insert polynucleotide seqid 32608.

XX

KW plant protectant; plant growth regulant; gene therapy; plant;

KW recombinant DNA construct; physical array; plant breeding marker;

KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;

KW extreme osmotic condition; pathogen tolerance; pest tolerance;

KW growth rate; cell cycle pathway; disease resistance;

KW galactomannan production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;

KW protein content; gene; ss.

XX

OS Unidentified.

XX

PN US2004034888-A1.

XX

PD 19-FEB-2004.

XX

PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAJ/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
FI WPI; 2004-180133/17.
DR
XX
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX
PS Claim 1; SEQ ID NO 32608; 15pp; English.
XX
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX
SQ Sequence 1804 BP; 464 A; 392 C; 472 G; 476 T; 0 U; 0 Other;

Query Match 7.9%; Score 195.8; DB 13; Length 1804;
Best Local Similarity 63.2%; Pred. No. 1.6e-42;
Matches 371; Conservative 0; Mismatches 127; Indels 89; Gaps 1;

QY 647 GCGAGGCGTAGAGCGGTGGTCCATCGCTTCACAGAGGATTTGCTGTGATTGTGGA 706
DB 394 GGTGAGACGACAGAGCTGTTGTGCTTCCCTGCACAAAGATTTCTGTCCATTATCGTGA 453

QY 707 GAGAGTATTCTGCAATGTGCAGAGAGGCTTGTGAGAGAGTTGTCAGCCTTATGCAGAAGA 766
DB 454 CRAAGGTATTGTAATGCTGAGAGACTATAGACAACACTGAGCCATATGCTTTGAG 513

QY 767 CGGAAGTGTGTTGAATATGGAACGAGGTTCTCTCAGATGACATTTGATGTCAATGGGTT 826
DB 514 TGGGGAACTGTCAATATGGAACGAGGTTTCTCAGTTGACATTTGATGTGATTTGTTT 573

QY 827 GTCTCTTTTAACTACATTTTCATTTCTTTGACTACTGATGCTTCCTGATTTGAACTGT 886
DB 574 ATCATTTGTTCACTACATTTTGAATTCCTTCCCTCAACACAGATAGTCTGTCAATGATGCTGT 633

QY 887 TTACATGCTCTTAAAGAGCTGAGCTTCCTTCTGATCTTCTGCAATATTTGGAAGGC 946
DB 634 TTATATGCACTCAAGAGACAGAGCTTCCTGCTCAGATCTTTGCACTATGGAAGGT 693

QY 947 AAGTTTCCTGTGTTTCTTCTGTGGTTTGTGTTGATTGTGTGGAACAATTTGGAATCTTTGTTAA 1006
DB 694 TGGTTTC----- 700

QY 1007 TTGAGAGGGTTTGGTTGTTTCTTTTTCAGATCGATGCTATTGTTAGATAGTCCCGAGACAG 1066

Db 701 -----TTGTCAAGATAATCCCAAGACAG 724

QY 1067 GTGAAGCTGAAGAGCTGTAACTTTTGTATAGGGAACCTGTTGAAGACCTTTATTGCTTAAG 1126
Db 725 ATAAAGACGAGAGAAATGCGGTTTACGATTATATAAGGAACACTGTTTGAAGAGCTGATTATGAAG 784

QY 1127 TGTAAAGAAATTTGCGAAAGAGAGGCGGAAGAAGAAATCAATGATGAGGAGTATGTTAAATGAT 1186
Db 785 TGTAAAGAAATAGTGGAGAGCTGAAATGACAGATTGAGGCTGAGGATATATGTTAAACGAA 844

QY 1187 GCTGACCCCAAGTATCTGCGTTTCTTGTCTTCAAGCAGAGAGAGGT 1233
Db 845 GGGGATCCTAGCATTCTACGCTTCTCTACTTGTCTAGCCGAGATGAGGT 891

RESULT 6
AEB16926
ID AEB16926 standard; cDNA; 1086 BP.
XX
AC AEB16926;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 24.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Triticum aestivum.
XX
FH Location/Qualifiers
FT 1..1086
FT /*tag= a
FT /product= "wheat cytochrome P450 97C (CYP97C) protein"
FT /partial
FT /note= "No stop codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
DR P-PSDB; AEB16920.
DR GENBANK; CA497665, BG906289, CA742365, CA742792.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 24; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein-epoxide hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC (CYP97C) cDNA.
XX
SQ Sequence 1086 BP; 243 A; 298 C; 295 G; 250 T; 0 U; 0 Other;

```
Query Match      7.6%; Score 187.2; DB 14; Length 1086;
Best Local Similarity 61.9%; Pred. No. 2.9e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;

QY 647 GCGAGGCGTAGAGCGGTGCTCCATCGCTTCACAGGAGTATTTGTCTGTGATTTGCGA 706
DB 489 GGTGAGACGTAGACGCGTGTACCATCTCTACACAAAGATTTCTCTGATTAATGTCGA 548

QY 707 GAGAGTATTTCTCAATGTCAGAGAGCGCTTGTGTGAGAAAGTTGCGAGCTTTATCGAAGA 766
DB 549 TAAAGTGTCTGATAATGTCTGAGAGATTTGTGTGAAAGCTCGAGACTTATGCTTTGAG 608

QY 767 CGGAGTGTCTGATAATGTCAGAGCGAAGTTCTCTGATGACATGATGATGATGATGATG 826
DB 609 TGGTGAACCTGTAAATATGGAAGCGAGGTTTCTCAATGACATGATGATGATGATGATG 668

QY 827 GTCTCTTTTAACTCAATTTCTGATTTCTTGTGACTGATGATGATGATGATGATGATG 886
DB 669 ATCTTGTTCACACTTGAATTCCTCTCAGATGATGATGATGATGATGATGATGATG 728

QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 946
DB 729 TTACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 783

QY 947 AAGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTTCT 1006
DB 784 ----- 783

QY 1007 TTGAGAGGTTTGGTTGTTTCTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGAC 1066
DB 784 -----CAGATCGAATTTGCTGTGCAAGATTTGTTCTTAGACAG 819

QY 1067 GTGAAAGCTGAAAGCGCTGTAACTTTGATAAGGGAACCTGTTGGAAGACCTTTATGCTAAG 1126
DB 820 ATAAAGCGGAAAGAGCAGTTAAACATTAAGGATACCGTTGAGAGCTTAATTACAAA 879

QY 1127 TGTAAAGAAATTTGTCGAAGAGAGCGGCAAGAAATCAATGATGAGAGATGATTAATGAT 1186
DB 880 TCGAAGGCAATCGTAGATCTGAAATGAACAGATTTGAGGGTGAAGAAATATGTAATGAG 939

QY 1187 GCTGACCAAGTATCTCGTTTCTTCTGCTTCAACAGAGAGAGGTTTAACTTT 1242
DB 940 GCAGATCCTAGCATCCTCGGTTTCTTACTTGTAGCCGTTGAAGAGGTTCAGCAGTTT 995

RESULT 7
ID ABB16925 standard; cDNA; 1638 BP.
XX AC AEB16925;
XX DT 08-SEP-2005 (first entry)
XX DE Barley cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 23.
XX KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
XX OS cytochrome P450 97C; gene; ss.
XX OS Hordeum vulgare; subsp. vulgare.
XX OS Hordeum vulgare; subsp. spontaneum.
XX FH Key
XX CDS Location/Qualifiers
FT 1..1638
FT /*tag= a
FT /product= "Barley cytochrome P450 97C (CYP97C) protein"
FT /transl_except= (pos:1585..1587, aa:Xaa)
FT /note= "Xaa may be any naturally occurring amino acid"
XX PN US2005150002-A1.
XX XX
XX PD 07-JUL-2005.
XX XX
```

PP 02-JAN-2004; 2004US-00751235.
XX 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIM/) KIM J.
PI Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.
DR P-PSDB; AEB16919.
DR GENBANK; EMB16653, BU987393, CA023004.
DR DDBJ; AV835803.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 23; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using luteal epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX
SQ Sequence 1638 BP; 406 A; 403 C; 427 G; 401 T; 0 U; 1 Other;
Query Match 7.6%; Score 187.2; DB 14; Length 1638;
Best Local Similarity 61.9%; Pred. No. 3.5e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;
QY 647 GCGAGGCGTAGAGCGGTGCTCCATCGCTTCACAGGAGTATTTGTCTGTGATTTGCGA 706
DB 474 GGTGAGACGTAGACGCGTGTACCATCTCTACACAAAGATTTCTCTCAGTAATGCTTGA 533
QY 707 GAGAGTATTTCTCAATGTCAGAGAGCGCTTGTGAGAAAGTTGCGAGCTTTATGCGAAGA 766
DB 534 TAAAGTGTGTTGTAATGTCGAGAGATTTGTGGAAGCTCGAGACATATGCTTTGAG 593
QY 767 CGGAGTGTCTGATAATGTCAGAGCGAAGTTCTCTCAGATGACATCTGATGATGATGATG 826
DB 594 CGGTGAACCTGTTAATATGGAAGCGAGATTTTCTCAATGACACTAGATGATGATGATG 653
QY 827 GTCTCTTTTAACTCAATTTGATTTCTTGTGATCTGATGATGATGATGATGATGATG 886
DB 654 GTCTTTGTTCAACTACAACTTTGATTTCTCTCAGATGATGATGATGATGATGATG 713
QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 946
DB 714 TTACCCGCACTGAAAGAGAGAGGCTCGTTTCTCAGATCTTTTACCATCTG----- 768
QY 947 AAGTTTCTGTGTTTCTTCTGTTTCTGATTTGTGTTGGAACAATTCGATTTCTGTTAA 1006
DB 769 ----- 768
QY 1007 TTGAGAGGTTTGGTTGTTTCTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
DB 769 -----CAGATTTGCTGTGCAAGATTTGTTCTTAGACAG 804
QY 1067 GTGAAAGCTGAAAGCGCTGTAACTTTGATAGGGAACCTGTTGAGAGACCTTTATGCTAAG 1126
DB 805 ATCAAGCAGAAAGGCGAGTTTAAACAATAAGGATGATGTTGTAAGAGCTTAATTAA 864
QY 1127 TGTAAAGAAATTTGTGGAAGAGAGAGCGGCAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
DB 865 TGCAGGCAATCGTAGATGCTGTAATAAGAACAGATTTGAGGGTGAAGAAATATGTAATGAG 924


```
PR 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16938.
DR GENBANK; BF425906, BF596805, AW704660, AW704625, BI470164, BQ296458,
DR BMS92469, AI938600, AI938382, BU544173, BI471346, CD410775, BF598710,
DR BG154747.
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX Claim 10; SEQ ID NO 45; 135pp; English.
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using IUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is soybean cytochrome P450
CC monooxygenase (CYP97A) cDNA.
XX Sequence 1278 BP; 389 A; 247 C; 297 G; 345 T; 0 U; 0 Other;
SQ
Query Match 3.7%; Score 90.8; DB 14; Length 1278;
Best Local Similarity 59.2%; Pred. No. 7.8e-14;
Matches 174; Conservative 0; Mismatches 117; Indels 3; Gaps 1;
QY 651 AGCGGTAGAGCGGTGTCATCGCTTACAGGAGGTATTTCTGTGATTTGTGGAGAGA 710
Db 43 AGACGTCGTGTATAGTCCAGCATTTGCCACCAAGATGATGTAGCAGCTATGATTG---GC 99
QY 711 GTATTCTGCATATGTCAGAGAGGCTGTTGAGAAAGTTCAGCCCTTATGCAGAACGGA 770
Db 100 CTTTTCGGAAGAAGTCAGATAGGCTCTCCAGAACTAGATGCTGTCATCCGATGA 159
QY 771 AGTGCTGTCAATATGAAGCGAAGTTCTCTCAGATGACACTTGTATGTCATGCGTGTCT 830
Db 160 GAAGATGTTGAGATGGAATCACTTTTCTCTCGATTGACCTTGGACATCATTTGCCAAGCA 219
QY 831 CTTTTTAACTACAAATTCGATTTCTTGACTACTGATAGTCTCTCATTTGAAGCTGTTTAC 890
Db 220 GTATTCAATATGATTTTGATAGTTTATCAAAATGACACCTGGTATAGTTGAGGCTGTTTAT 279
QY 891 ACTGCTCTTAAGAAGCTGAGCTTCGTTCTTCTACTGATCTTCTGCCATATTGGAAG 944
Db 280 ACTGTACTGAGAAAGCAGAAAGTCAAGTGTGCTCCAAATCCAGTCTGGGAG 333
RESULT 15
AD082389 standard; cDNA; 1078 BP.
XX AC
XX AD082389;
XX 21-APR-2005 (first entry)
XX Plant full length insert polynucleotide seqid 1109.
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
```

```
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX Unidentified.
XX US2004034888-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 1109; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX Sequence 1078 BP; 332 A; 180 C; 253 G; 313 T; 0 U; 0 Other;
SQ
Query Match 3.5%; Score 87.4; DB 13; Length 1078;
Best Local Similarity 69.8%; Pred. No. 6.1e-13;
Matches 118; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 2177 CAGATTTCATCCATTTCAGTGGAGGCGCTAGAAATGTGTAGCGATCAGTTTGATTGAT 2236
Db 186 CAGGTTTATTCATTTCAGTGGGGGCTCTCGGAAATGTGTGGAGATCAATTCCTCTCCT 245
QY 2237 GGAGGCAATTGTGGCACTCGCGTGTCTTTCAGCGGTTTAAAGCTTGAGCTGTTCTCTGA 2296
Db 246 AGAAGCAATAGTGGCACTTGCAGTTGTGTTCAGAAGATGATATCAGCTTGTGCCAGA 305
QY 2297 TCAGACCATTTAGCATGACCACAGGAGCAACCATATACACACCAATGTA 2345
Db 306 TCAAAAGATTAAACATGACTACTGAGCTACAAATTCATACACACAGTGA 354
Search completed: January 13, 2006, 01:04:32
Job time : 1451 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 23:46:35 ; Search time 9655 Seconds
(without alignments)
11954.814 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagttctactcttttc.....tgaaggtgagccaaaggttaa 2467

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791	32.1	1709	4	CNS0A665
2	723.2	29.3	1664	4	Arabidopsis
3	701.4	28.4	736	10	Arabidopsis
4	499	20.2	616	1	CW839414 GT8284.Ds
5	473	19.2	977	10	AV827314 AV827314
6	442	17.9	657	10	CL469740 SAIL 133
7	429.4	17.4	708	9	CB255014 54-E01836
8	386.6	15.7	570	9	BH931006 odillog10
9	323	13.1	559	6	BH584135 BOHAD90TF
10	323	13.1	716	6	CD828034 BN25.0691
11	322.8	13.1	664	6	CX189683 72-E02430
12	312.4	12.7	627	6	CD828321 BN25.070G
13	301.2	12.2	310	10	CD828299 BN25.070F
14	228.6	9.3	858	7	BX531442 Arabidopsis
15	220.4	8.9	704	10	CO072553 GR_Eas1M
16	218.4	8.9	247	10	CZ710414 OC_Ba003
17	211.8	8.6	897	8	CW839413 GT8284.Ds
18	211	8.6	914	7	CV886111 UCRC04 2
19	209.6	8.5	563	2	CK290930 EST753644
20	204.8	8.3	886	8	BF096935 EST400472
21	203.2	8.2	604	8	DR923995 EST111553
22	203	8.2	386	3	CX186676 D11 45-75
					BP827113 BP827113

23	202.2	8.2	642	8	CX176509
24	199.8	8.1	668	8	DR952300
25	194.2	7.9	447	1	DR952300 EST114383
26	192.6	7.8	721	7	AV422798
27	191.8	7.8	469	8	CV529617 ALV 011A
28	191.6	7.8	650	10	CX541939 b13DNF0BH
29	191.6	7.8	652	2	CZ799589 OC_Ba017
30	187.2	7.6	600	5	BI263819 NF114G06P
31	187	7.6	693	10	BU987393 HF14K04r
32	185.6	7.5	722	8	CW166683 104 575 1
33	185	7.5	444	1	DR403827 CSAH-PNPI
34	185	7.5	511	2	AI736606 sb31b06.Y
35	185	7.5	530	3	BF597405 su97408.Y
36	182.4	7.4	589	7	BI701721 sail19d02.
37	181.6	7.4	440	5	CN182757 UCRC04 0
38	180.8	7.3	480	1	BQ585965 E012531-0
39	180.8	7.3	596	1	AJ486563 AJ486563
40	179.2	7.3	733	7	AJ432371 AJ432371
41	178.4	7.2	451	8	CV712131 UCRP01 0
42	175	7.1	689	5	CX023647 Md1v4-404
43	173.8	7.0	941	4	BQ862275 QGC20K13
44	170.4	6.9	457	1	BX826133 Arabidops
45	169.6	6.9	626	5	AW617348 EST323759
					BQ971938 QHB9C20.Y

ALIGNMENTS

RESULT 1	CNS0A665	1709 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB64ZA07 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX822785				
VERSION	BX822785.1	GI:42464524			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1709)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencing: BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life technologies (a division of Invitrogen) members carried out full-length libraries construction: Temple G. Genoscope members carried out sequencing and annotation: Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
source	1..1709				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				


```
QY 1438 TTCTACTTTGCTATGCTAGAGATTTACTTGGATGGAGCTTCTCTGTTCTCAATTACCT 1497
Db 1121 -----
QY 1498 CTTCAAAATTCCTATGTTCAVAGAACTCATCTGCAATTAAGGAAACACACAGAAGAAGTAG 1557
Db 1121 -----AGAACTCTCTGCAATTAAGGAAACACACAGAAGAAGTAG 1159
QY 1558 ACAGAGTGTAAAGAAAGAAACCCGGCTTTTCAGAGATATAAGAGATTGAATCATCA 1617
Db 1160 ACAGAGTGTAAAGAAAGAAACCCGGCTTTTCAGAGATATAAGAGATTGAATCATCA 1219
QY 1618 CTCGTTGTATAACGAGTCAATGCGTCTCTATCCCTCATCCCTCTGTAAGCAATCAAGCTC 1677
Db 1220 CTCGTTGTATAACGAGTCAATGCGTCTCTATCCCTCATCCCTCTGTAAGCAATCAAGCTC 1263
QY 1678 ATCTCTCTAAATTAATTCATGAACATAAATTTCTGATTTGATTTGTTTCTCGTAGGTCTTGA 1737
Db 1264 -----GTCTTGA 1270
QY 1738 TAAGAGAGCTCAAGTTCCTGACATTCCTGGAACATATAAGGTCAATACCGGACAAG 1797
Db 1271 TAAGAGAGCTCAAGTTCCTGACATTCCTGGAACATATAAGGTCAATACCGGACAAG 1330
QY 1798 ACATTTATGATTTACGTTCTATACATCCATCGTTCTTCGAGGTA 1841
Db 1331 ACATTTATGATTTACGTTCTATACATCCATCGTTCTTCGAGGTA 1374
```

RESULT 3

```
CW839414 736 bp DNA linear GSS 26-NOV-2004
LOCUS GT8284.Ds5.11.07.2002.jw72.736 Arabidopsis thaliana Landsberg Ds
DEFINITION insertion lines Arabidopsis thaliana genomic clone GT8284, genomic
survey sequence.
```

ACCESSION

```
CW839414
```

VERSION

```
CW839414.1 GI:56095204
```

KEYWORDS

```
GSS.
```

SOURCE

```
Arabidopsis thaliana (thale cress)
```

ORGANISM

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
```

```
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
1 (bases 1 to 736)
```

```
May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
```

```
McCombie,W.R. and Martienssen,R.A.
```

```
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
```

```
in transgenic lines
```

```
Unpublished (2004)
```

```
Contact: Martienssen RA
```

```
Cold Spring Harbor Laboratory
```

```
1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA
```

```
Tel: 516 367 8322
```

```
Fax: 516 367 8369
```

```
Email: martiens@cshl.org
```

```
This sequence flanks a Ds transposon carrying a gene trap in line
```

```
GT8284. The transposon is located within At3g53130.
```

```
Class: transposon-tagged.
```

FEATURES

```
Location/Qualifiers
```

```
1..736
```

```
/organism="Arabidopsis thaliana"
```

```
/mol_type="genomic DNA"
```

```
/scotype="Landsberg erecta"
```

```
/db_xref="taxon:3702"
```

```
/clone="GT8284"
```

```
/lines="Arabidopsis thaliana Landsberg Ds insertion
```

```
lines"
```

```
/notes="Lines of Arabidopsis thaliana were generated which
```

```
each contain a Ds transposon carrying a glucuronidase
```

```
reporter gene. Genomic DNA flanking the transposon
```

```
insertion in each line was amplified by PCR and directly
```

```
sequenced. More information is available at
```

```
http://genetrapp.cshl.edu."
```

ORIGIN

```
Query Match 28.4%; Score 701.4; DB 10; Length 736;
Best Local Similarity 98.5%; Pred. No. 9.5e-177;
Matches 717; Conservative 0; Mismatches 9; Indels 2; Gaps 1;
```

```
QY 274 CTCCTCGGAGGTGCTCTCTCTTACCTCTCTCAAAATGGATGAATGAGTAGCGACCCATT 333
Db 9 CTGCTCGGAGGTGCTCTCTTCTTACCTCTCTCAAAATGGATGAATGAGTAGCGACCCATT 68
QY 334 TACCGTCTCGCTGCTGCTCGTAATTTCTGTAATTTGTGAGGACCCACGAGTAGCTAAA 393
Db 69 TACCGTCTCGCTGCTGCTCGTAATTTCTGTAATTTGTGAGGACCCACGAGTAGCTAAA 128
QY 394 CATGTTTTGAGGAATTAATCCAAAGTACGCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTT 453
Db 129 CATGTTTTGAGGAATTAATCCAAAGTACGCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTT 188
QY 454 CTATTTGGTTCGGGTTTCGCTATCGCTGAAGGACCTCTTTGGACAGTAATTTTCATCTCT 513
Db 189 CTATTTGGTTCGGGTTTCGCTATCGCTGAAGGACCTCTTTGGACAGTAATTTTCCTCC 248
QY 514 CCT--ATCTCAATTTTGAAGTTTTTGGAAATTTGTGGAAGTAATGTGTGACTGTCTGTATG 571
Db 249 CCTCGACCACCAANNNTGAAGTTTTTGGAAATTTGTGGAAGTAATGTGTGACTGTCTGTATG 308
QY 572 ATAGTAACTCTAAATTTTAGGGTTTGTAGATTCCTCAATCTCTCTATTGGGCTTAGCTGAACT 631
Db 309 ATAGTAACTCTAAATTTTAGGGTTTGTAGATTCCTCAATCTCTCTATTGGGCTTAGCTGAACT 368
QY 632 CTGATTTTTCATAGGCGAGCGGTAGAGCGGTGTTCCATCGCTTCAACAGGAGTATTTT 691
Db 369 CTGATTTTTCATAGGCGAGCGGTAGAGCGGTGTTCCATCGCTTCAACAGGAGTATTTT 428
QY 692 GTCTGTGATTTGTGAGAGAGTATTTCTGCAAAATTCGACAGAGGCTTGTGTGAGAGTTGCA 751
Db 429 GTCTGTGATTTGTGAGAGAGTATTTCTGCAAAATTCGACAGAGGCTTGTGTGAGAGTTGCA 488
QY 752 GCCTTATCGCAAGGACGGAAGTGTGTAATGAAGCGAAGTTCTCTCAGATGACACT 811
Db 489 GCCTTATCGCAAGGACGGAAGTGTGTAATGAAGCGAAGTTCTCTCAGATGACACT 548
QY 812 TGATGCTCAATTTGGTGTGCTCTTTTTTAACTCAAAATTTTCGATTTCTTGACTACTGATGCTC 871
Db 549 TGATGCTCAATTTGGTGTGCTCTTTTTTAACTCAAAATTTTCGATTTCTTGACTACTGATGCTC 608
QY 872 TGTCATGAAGCTGTTTACATGCTCTTTAAAGAGCTGAGCTTCTTCTACTGATCTTCT 931
Db 609 TGTCATGAAGCTGTTTACATGCTCTTTAAAGAGCTGAGCTTCTTCTACTGATCTTCT 668
QY 932 GCATATTTGGAGGCAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 991
Db 669 GCATATTTGGAGGCAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 728
QY 992 TTGGATTC 999
Db 729 TTGGATTC 736
```

RESULT 4

```
AV827314
```

```
LOCUS AV827314
```

```
DEFINITION AV827314 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-13-p07 5',
```

```
mRNA sequence.
```

```
AV827314
```

```
ACCESSION AV827314.1 GI:19869374
```

```
VERSION AV827314
```

```
KEYWORDS EST.
```

```
SOURCE Arabidopsis thaliana (thale cress)
```

```
ORGANISM Arabidopsis thaliana
```

```
SpERMATOPHYTES; Magnoliophyta; eudicotyledons; core eudicotyledons;
```

```
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
1 (bases 1 to 616)
```

```
REFERENCE
```

```
1 (bases 1 to 616)
```



```

Db      297  AAAGCTGAGGAATTTCTCCCTGAACGATTCGACATAGATGGGCAATCCCTAACGAACA 238
Qy      2069  AACACTGATTTCAAGTAAACTCAGTAGAACACATCTTTTGTACACAAACTACTGAATCAAG 2128
Db      237  AACACTGATTTCAAGTAAACTCAGTAGAACACATCTTTTGTACACAAACTACTGAATCAAG 178
Qy      2129  ATTAGTGGTTTGTAGTAGGGAATTTAAAGATGATTTCTTTTTCACCAAGATTCATCCC 2188
Db      177  ATTAGTGGTTTGTAGTAGGGAATTTAAAGAGATTTTCTTTTTCACCAAGATTCATCCC 118
Qy      2189  ATTCACTGGAGGCGCTAGAAA 2210
Db      117  ATTCACTGGAGGCGCTAGAAA 96

RESULT 6
LOCUS   CB255014
DEFINITION 54-E018363-019-007-I14-T7R MP1Z-ADIS-019 Arabidopsis thaliana cDNA
clone MP1Zp768L147Q 5-PRIME, mRNA sequence.
ACCESSION CB255014
VERSION   CB255014.1 GI:56915639
KEYWORDS EST.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 657)
REFERENCE
AUTHORS Jakoby,M., Stracke,R., Soerensen,T.R. and Weisshaar,B.
TITLE Arabidopsis thaliana cDNA library enriched in transcription factors
JOURNAL Unpublished (2003)
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@piz-koeln.mpg.de
Insert Length: 657 Std Error: 0.00
Plate: 7 row: L column: 14
Seq primer: T7R; CTATACGACTCACTATAGGA.

FEATURES
source
1..657
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="At7"
/db_xref="GABI:597448"
/db_xref="taxon:3702"
/clone="MP1Zp768L147Q"
/tissue_type="hypocotyl"
/dev_stage="tissue culture"
/lab_host="E. coli DH5alpha mcr"
/clone_lib="MP1Z-ADIS-019"
/notes="Vector: pSPORT1; Site_1: NotI primer adapter;
Site_2: SalI primer adapter; RNA from cellculture (At7) 5
days after inoculation treated with 0.002 mM cycloheximid
for 2 h in the dark. Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de. This
clone is available from RZPD; contact RZPD (clone@rzpd.de)
for further information."

ORIGIN
Query Match 17.9%; Score 442; DB 6; Length 657;
Best Local Similarity 79.5%; Pred. No. 4.8e-107;
Matches 630; Conservative 0; Mismatches 10; Indels 152; Gaps 3;

Qy      155  AATCCCAATCATGGGTCAAGTCCGATGGCTCAACACTCACTCTGACCTTCTCTCAG 214
Db      5  AATCCCAATCATGGGTCAAGTCCGATGGCTCAACACTCACTCTGACCTTCTCTCTT 63
Qy      215  GAAAAACGACGATCAGGTATACCAATCCGAACCGAGCTCGACGATGTCGCTGATC 274
Db      64  GACANAACGACGATCAGGTATACCAATCCGAACCGAGCTCGACGATGTCGCTGATC 123

```

```

Qy      275  TCCTCGGAGGTCTCTCTTCTTACCTCTCTACAAATGGATGAATGAGTACGAGCCCATTT 334
Db      124  TCCTCGGAGGTCTCTCTTCTTACCTCTCTACAAATGGATGAATGAGTACGAGCCCATTT 183
Qy      335  ACCGTCCTCGTCTGGTCTCTGTAATTTGTAATTTGTAAGCGACCCAGCGATAGCTAAAC 394
Db      184  ACCGTCCTCGTCTGGTCTCTGTAATTTGTAATTTGTAAGCGACCCAGCGATAGCTAAAC 243
Qy      395  ATGTTTGTAGGAATTAATCCAAAGTACGCTAAAGCGTTAGTTCGCTGAAAGTCTCTGAAATTC 454
Db      244  ATGTTTGTAGGAATTAATCCAAAGTACGCTAAAGCGTTAGTTCGCTGAAAGTCTCTGAAATTC 303
Qy      455  TATTTGGTTTCGGGTTTCGCTATCGCTGAAGCACTCTTTGGACAGTAAATTTTCATCTCCTC 514
Db      304  TATTTGGTTTCGGGTTTCGCTATCGCTGAAGCACTCTTTGGACA----- 347
Qy      515  CTATCTCAATTTTGAAGTTTTTGGAAATTTGTGGAAGTAAATGTGTGACTCTCTTTGTATGATA 574
Db      348  ----- 347
Qy      575  AGTAACCTCTAAATTTTAGGGTTTAGATTCCAATCTCTCTATTTGGGCTTAGCTGAAGTCTG 634
Db      348  ----- 347
Qy      635  ATTTTTCATATAGGCGAGCGGTAGAGCGGTGTTCCATCGCTTTCACAGGAGGTATTTGTC 694
Db      348  -----GCGAGCGGTAGAGCGGTGTTCCATCGCTTTCACAGGAGGTATTTGTC 394
Qy      695  TGTGATTTGTGAGAGAGTATTTCTGCAATGTGACAGAGGCTTGTGTGAGAAAGTTGAGAGCC 754
Db      395  TGTGATTTGTGAGAGAGTATTTCTGCAATGTGACAGAGGCTTGTGTGAGAAAGTTGAGAGCC 454
Qy      755  TTATCGAAGACCGGAAGTCTGTGTAATATGGAACGGAAGTTCTCTCAGATGACACTTGA 814
Db      455  TTATCGAAGACCGGAAGTCTGTGTAATATGGAACGGAAGTTCTCTCAGATGACACTTGA 514
Qy      815  TGTCAATCGGTTGCTCTCTTTTAACTCAATTTCCGATCTTTTGAATCTACTGATGAGTCTCT 874
Db      515  TGTCAATCGGTTGCTCTCTTTTAACTCAATTTCCGATCTTTTGAATCTACTGATGAGTCTCT 574
Qy      875  CATTTGAAGCTGTTTACACTGCTCTTTAAAGAA--GCTGAGCTTCGTTTACTGATCTTCTG 932
Db      575  CATTTGAAGCTGTTTACACTGCTCTTTAAAGAAAGCTTGAGCTTCGTTTACTGATCTTCTG 634
Qy      933  CCATATTGGAAG 944
Db      635  CCATATTGGAAG 646

RESULT 7
BH931006/c
LOCUS BH931006/c
DEFINITION odii10g10.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BH931006
VERSION BH931006.1 GI:23411072
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 708)
REFERENCE Delchaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
AUTHORS Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: odii10 row: g column: 10

```


Seq primer: -28RPpOT reverse
 Class: shotgun
 High quality sequence start: 81
 High quality sequence stop: 434.
 Location/Qualifiers
 1. .708
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.Oleracea002"
 /notes="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 17.4%; Score 429.4; DB 9; Length 708;
 Best Local Similarity 81.9%; Pred. No. 1.2e-103;
 Matches 525; Conservative 0; Mismatches 101; Indels 15; Gaps 2;
 QY 626 TGAAGTCTGATTTTACATAGCGAGCGGTAGAGCGGTGTTCCATCGCTTCACAGGAG 685
 Db 672 TGAATTCCTGTTGATCTTAGCGAGGAGAGCGGTGTTCCGTCACCTCACAAGAG 613
 QY 686 GTATTGCTGTCATTTGCGAGAGATATCTGCAATATGTCAGAGAGCGCTTGTGAGAA 745
 Db 612 GTACTTGTCTGATTTGCGAGAGATGTTCTGCAGATGTCAGAGAGCGCTGTTGAGAA 553
 QY 746 GTTGCAGCGCTTATGCAAGAGACGGAAGTCTGTGAATATGGAAGCGAAGTCTCTCAGAT 805
 Db 552 GCTGCAGCGCTTATGCAAGTGCAGGAGAAAGCTGTGAACATGGAAGAGAACTCTCAGTT 493
 QY 806 GACACTTGATGTCATTTGGCTCTCTTTTAACTACAAATTCGATTCCTTTCAGTACTCA 865
 Db 492 AACTCTGCAGCGATTTGGCTTATCTCTTTTAACTGCAACTTGTATTCACCTCAGTACCGA 433
 QY 866 TAGTCTGTCTATGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGA 925
 Db 432 TAGTCTGTCTATGCAAGCTGTTTACACTGCTCTCAAGAGAGCTGAGCTTCGTTCTACTGA 373
 QY 926 TCTTCTGCCATATGGAAGCGCAAGTTTCTGTGTTTTTCTGTGTTTGTGATTTGTTG 985
 Db 372 TATTTCTACCTTATGGAAGCGCAAGTTTGTGTTGTTTCTTCTGTG-----C 327
 QY 986 GAACAATGGATCTTGTATTTAGAGGGT-TTGGTTGTTTTTTTTCAGATCGATGCATT 1044
 Db 326 GCAGACTTGGATCTTGTGAGCGCTTTGGTTGTTTCTTGTTCAGATCGATGCGTT 267
 QY 1045 GTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAAGGCTGTAACTTTTGATAAGGAAAC 1104
 Db 266 GTGTAAGATAGTCCCGAGACAGGTGAAAGCGCAAGAGGCTGTGACTTTGATAAGGAAAC 207
 QY 1105 TGTGGAAGACCTTATTTGCTAAGTAAAGAAATGTCGAAAGAGAGCGCAAGAAATCAA 1164
 Db 206 AGTTGAGGATCTCATTTGCAAAATGCAAGAGATTTGGAAGAGAGCGCAAGAAATCAA 147
 QY 1165 TGATGAGGATATGTAATGATGCTGACCCAGATATCTGCGTTTCTTGTGCTTGCAGCAG 1224
 Db 146 TGATGAGGATATGTAATGATGCTGACCCAGATATCTTGCCTTCTTGTGCTTGCAGCAG 87
 QY 1225 AGAAGAGGTTTAAACTTTTCTTAAAGTTTATAGCAAT 1265
 Db 86 AGAAGAGGTTTAACTTTCTTTTATATTCATCCATAT 46

RESULT 8

BH584135/c 570 bp DNA linear GSS 15-DEC-2001
 LOCUS BH584135
 DEFINITION BOHAD90TF BOHA Brassica oleracea genomic clone BOHAD90, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BH584135
 BH584135.1 GI:17836592
 GSS.
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
 AUTHORS

1 (bases 1 to 570)
 Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
 Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its
 application to gene discovery and annotation in Arabidopsis
 Genome Res. 15 (4), 487-495 (2005)

JOURNAL
 PUBMED
 COMMENT

15805490
 Other GSSs: BOHAD90TR
 Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TP

Class: sheared ends.

FEATURES
 Location/Qualifiers

1. .570
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone_lib="BOHAD90"
 /note="Vector: PHOS1; Site 1: BatX1; 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BatX1 linkers"

ORIGIN

Query Match 15.7%; Score 386.6; DB 9; Length 570;
 Best Local Similarity 82.0%; Pred. No. 3.6e-92;
 Matches 475; Conservative 0; Mismatches 89; Indels 15; Gaps 2;
 QY 624 GCTGAAGTCTGATTTTTCATAGCGAGCGGTAGAGCGGTGTTCCATCGCTTCACAGG 683
 Db 565 GTTGTATCTTGTGTTGATGTTAGCGAGGAGAGCGGTGTTCCGTCACCTCAAG 506
 QY 684 AGGTATTTGTCGTGATTTGTCAGAGAGATATCTGCAATATGTCAGAGAGCGCTTGTGAG 743
 Db 505 AGGTACTTGTCTGTGATTTGTCAGAGAGATGTTCTGCAGATGTCAGAGAGCGCTCGTTGAG 446
 QY 744 AAGTTGCAGCGCTTATGCAAGAGACGGAAGTGTCTGAAATATGGAAGCGAAGTTCTCTCAG 803
 Db 445 AAGCTGCAGCGCTTATGCAAGTGGAGGCAAAAGCTGTGACATGGAAGAGAAAGTTCTCTCAG 386
 QY 804 ATGACACTTGATGTCATTTGGGTTGCTCTTTTAACTACAAATTCGATTCCTTGACTACT 863
 Db 385 TTAACCTCGAGCGTGAATGGCTTATCTCTTTTAACTACAACTTTGATTCACCTACACC 326
 QY 864 GATAGTCTGTCAATGAAGCTGTTTACACTGCTCTTTAAAGAGCTGAGCTTCGTTCTACT 923
 Db 325 GATAGTCTGTCAATGAAGCTGTTTACACTGCTCTCAAGAGAGCTGAGCTTCGTTCTACT 266
 QY 924 GATCTTCTGCCATATTTGAAGSCAAAGTTTCCCTGTGTTTTTTTCTGTGTTTGTGATG 983
 Db 265 GATATTTCTACCCCTATTTGAAGSCAAAGTTTGTGTTTCTTCTGTGTTGTC----- 218
 QY 984 TGAACAATATGATTCCTGTTAAATGAGAGGT-TTGGTTGTTTTTTTTCAGATCGATGCA 1042
 Db 217 --GCAGACTTGGATTTCTTTGAGCGCTTTGGTGTGTTTTTCTTGTGTTTCAGATCGATGCG 160
 QY 1043 TTGTCTAAGATAGTCCCGAGACAGCTGAAAGCTGAAAGCGCTGTAACCTTTTGATAAGGAA 1102
 Db 159 TTGTCTAAGATAGTCCCGAGACAGTGAAGAGCAAGTGAAGAGAGGCTGTGACTTTTGATAAGGAA 100

[illegible]

AUTHORS	Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.
TITLE	Global assembly of Cotton ESTs
JOURNAL	Unpublished (2004)
COMMENT	Contact: Rod A. Wing Arizona Genomics Institute The University of Arizona Forbes Building Room 303, Tucson, AZ, 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: rwing@genome.arizona.edu Plate: 31 Tow: M column: 14.
FEATURES	Location/Qualifiers
source	1..858 /organism="Gossypium raimondii" /mol_type="mRNA" /db_xref="taxon:29730" /clone="GR_Ea3IM14" /tissue_type="whole seedlings" /dev_stage="first true leaves" /lab_host="DH10B" /clone_lib="GR_Ea" /note="Vector: pCMV_SPORT-6.1; Site.1: NotI; Site.2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."
ORIGIN	
Query Match	9.3%; Score 228.6; DB 7; Length 858;
Best Local Similarity	66.5%; Pred. No. 1.2e-49;
Matches 399; Conservative	0; Mismatches 109; Indels 89; Gaps 1;
Qy	652 GGCCTAGACGGGTGTTCCATCGCTTCACAGGAGGTATTTCTCTGTCGATTTGTGGAGAG 711
Db	1 GCGCGAGGGCTGTAGTTCATCTCTTCAATAAGTATTTGTCGTGTTATGTTGATCGG 60
Qy	712 TATTTGCAATGTGCAGAGGCTGTTTGGAAAGTTGCAGCCTTATGCAGAGACGGAA 771
Db	61 TATTTTGCAAATGCGCAGAGATTAGTAGAGAACTGCAACCTTTTGCATTAGATGCGCA 120
Qy	772 GTGCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGTCATTTGGGTGTCTC 831
Db	121 CTGCGGTGACATGGAAGAAAGTTTCTCAACTGACTCTTGATGTTATCGTCTCTCG 180
Qy	832 TTTTAACTACAATTCGATCTCTTTGACTACTGATGTCCTGTCAATGAAAGCTGTTTACA 891
Db	181 TATTTAATTATAACTTCGATTCGTTGACAACCGATAGCCCTGTCAATTGATGCAGTTTATA 240
Qy	892 CTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTCGAAGGCAAGTT 951
Db	241 CTGCTTTGAAAGAAGCAAGAAATTACGATCCACGGACATTTTACCCATATTGGA----- 291
Qy	952 TCCTGTGTTTTTCTGTGGTTTGTGTTGATGTGTGGACAATTCGATCTTTGTAATTGAG 1011
Db	292 ----- 291
Qy	1012 AGGGTTTGGTTGTTTTTTTCAGATCGATTCGATGTTGTAAGATAGTCCCGAGACAGGTGAA 1071
Db	292 -----AGATTAGTGTCTTTGTGCAAGATAGTTCGAGACAATAAAA 331
Qy	1072 AGCTGAAGAAGCTGTAACTTTGATAAGGAAA CTGTTTGAGACCTTATTCCTAAGTGTA 1131
Db	332 GGCTGAAAAAGCAGTTTACAGTCATTCGGAAGCCGTTTGAAGAACTTATTTGAAGTGCNA 391
Qy	1132 AGAAATGTGCAAGAGAAGCGCAAGAAATCAATGATGAGGAGTATGTAATGATGCTGA 1191
Db	392 AGAGATGTTGAAAAAGAGGTGNAAGGATCAACGAGAGGGAATATGTAATGATGCTGA 451
Qy	1192 CCCAAGTATCCTGCGTTTCTTGCTTGCAAGCAGAGAAAGAGGTTTAAACTTT 1242
Db	452 TCCAAGTATCCTTCGATCTTCTGTTGCAAGCAGGGAAGAGGTTTCAAGCTT 502

RESULT 15	CZ710414	704 bp	DNA	linear	GSS 25-JUL-2005				
LOCUS	OC_Ba0031117.r	OC_Ba	Oryza	coarctata genomic clone	OC_Ba0031117				
DEFINITION	Oryza coarctata genomic clone OC_Ba0031117								
ACCESSION	CZ710414	Genomic survey sequence.							
VERSION	CZ710414.1	GI:71116126							
KEYWORDS	GSS.								
SOURCE	Oryza coarctata (Porteresia coarctata)								
ORGANISM	Oryza coarctata								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.								
REFERENCE	1. (bases 1 to 704)								
AUTHORS	Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and Wing,R.								
TITLE	OMAP (Oryza Map Alignment Project) - Arizona Genomics Institute								
JOURNAL	Unpublished (2005)								
COMMENT	Contact: Rod A. Wing Arizona Genomics Institute University of Arizona Forbes Building Room 303, Tucson, AZ 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: rtwing@genome.arizona.edu								
	PCR Primers								
	FORWARD: TAA TAC GAC TCA CTA TAG GG BACKWARD: CAC TCA TTA GGC ACC CCA								
	Plate: 0031 row: 1 column: 17								
	Seq primer: CAC TCA TTA GGC ACC CCA								
	Class: BAC ends.								
FEATURES	Location/Qualifiers								
	1..704								
	/organism="Oryza coarctata"								
	/mol_type="genomic DNA"								
	/db_xref="taxon:77588"								
	/clone="OC_Ba0031117"								
	/tissue_type="leaves"								
	/dev_stage="mature"								
	/lab_host="DH10B"								
	/clone_lib="OC_Ba"								
	/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"								
ORIGIN									
	Query Match 8.9%; Score 220.4; DB 10; Length 704;								
	Best Local Similarity 64.8%; Pred.No.1.8e-47;								
	Matches 376; Conservative 0; Mismatches 171; Indels 33; Gaps 2;								
Qy	646	AGCGAGGGGTAGACGGGTGGTTCCATCGCTTCACAGGAGGATTTGTCTGTGATTTGG	705						
Db	124	AGGTGAGAGCGTCGATCAGTCGTACCATCTCTACACAAACGATTTCTCTCAGTGATGGTTG	183						
Qy	706	AGAGAGTATTCTGCAATGTGCAGAGAGGCTTCTTCAGAAAGTTGCAGCCCTTAGCAGAAG	765						
Db	184	ACAGAGTTTTTTGTAATGTGCTGAGAGATTAGTGAGAGAGCTTGAAGCTCTGCTTTAA	243						
Qy	766	ACGGAAAGTGTGTGAATATGGAGCGGAAGTTCTCTCAGATGACACTTGATGTCAATGGGT	825						
Db	244	GTGGCAAAACCTGTAAATATGGAGCAAGGTTCTCTCAATGACTTTTAGATGTGATGGTT	303						
Qy	826	TGTCCTCTTTTAACATAATTTGCATTTCTTTGACTACTGATAGTCTGTCTCATTTGAAGCTG	885						
Db	304	TGTCCTTGTTCATATTAATTTGATTTCCCTCACGTCAGATAGTCCAGTTATTGATGCTG	363						
Qy	886	TTTACACTGCTCTTAAGAAAGCTGAGCTTGGTTCTACTGATCTTCTGCCATATTTGGAAGG	945						
Db	364	TCTACACTGCACTCAGGAGAGAGAGCTTGGTTCTACAGATCTTTTACATACCTGGAGG	423						
Qy	946	CAAGTTTCCTGTGTTTTTTTCTGTGGTTGTGTA-----TTGTGTGGAAACAATTG	994						
Db	424	TATATTCCTTTTGCCAGTTTTTTTTTATAAAAAAATCTCATTCAGTTGGCTGTCTA	483						


```

; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1969
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-1969

Query Match      1.7%; Score 42; DB 7; Length 1923;
Best Local Similarity 52.9%; Pred. No. 0.25;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 2181 TTCATCCCAATTCAGTGGAGGCGCTAGAAAATGTGTAGCGGATCAGTTTGGCATTGATGGAG 2240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1329 TTCATGCGCTTTTCTGCGAGGCGCCAGGAATTCATCGGGCAGCATTTGCCATGAAACGAG 1388
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2241 GCAATTGTGGCACTCGCGGTGTTCTTCAGCGGTTAAACGTTGAGCTGGTTCTCGATCAG 2300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1389 ATGAAGTGGTCACAGCCCTTTGTTTACTGCGCTTTGAGTTCTCCCTGGATCCCTCAAAG 1448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2301 ACCATTAGCATGACCACAGGAGCAACCATACACACCAATGTATGCCA 2350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1449 ATGCCCAATTAAGTCCCCCAGCTGATCTTGGCGCTCCAAAATGGTATCCA 1498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-11-136-527-3880
; Sequence 3880, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3880
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3880

Query Match      1.6%; Score 40.6; DB 7; Length 2118;
Best Local Similarity 60.4%; Pred. No. 0.65;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 2181 TTCATCCCAATTCAGTGGAGGCGCTAGAAAATGTGTAGCGGATCAGTTTGGCATTGATGGAG 2240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1512 TTTATTTCCCTTCTCCGCTGGGCCAGGAACTGCATAGGACAGACTTTTGGCCATGAAACGAG 1571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2241 GCAATTGTGGCACTCGCGGTGTTCTTCAGCGGTTAAACGTTGAGCTGGTT 2291
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1572 ATGAAGTGGCACTGGCGCTGACGCTGCTGCGCTTCCGCGTCTCGCGGAT 1622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-11-121-086-12
; Sequence 12, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000

; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 12
; LENGTH: 119160
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-12

Query Match      1.6%; Score 40.2; DB 7; Length 119160;
Best Local Similarity 49.8%; Pred. No. 9.1;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 826 TGTCTCTTTTAACTACAAATTTGCAATTTCTTGACTACTAGTAGTCTCTGCTCATTTGAAGCTG 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95869 TGACAAATGTTTAACTGCAAAATTCATTTTCTTCAATAGTTTTAGGGATATTTGCTGTATTG 95928
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 886 TTACACCTGCTCTTAAAGAGCTGAGCTTCGTTCTTACTGATCTCTCTGCCATATTTGGAAGG 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95929 ATATAACTTCTCTTGTCTGATATTTGTAATTTTATATCTTTTCATTTTTCCTGAT 95988
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 946 CAAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95989 CAGTTGAGCTAGAGTTTAAATAATTTTATGATCTTTTGAAGGAACCAAGCTTTTGGTTT 96048
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1006 ATTGAGAGGGTTTGGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 1030
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96049 CATTGATTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 96073
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-11-136-527-5679
; Sequence 5679, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 5679
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5679

Query Match      1.6%; Score 39.4; DB 7; Length 600;
Best Local Similarity 43.1%; Pred. No. 0.67;
Matches 115; Conservative 13; Mismatches 139; Indels 0; Gaps 0;

QY 898 TTAAGAAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGCAAGTTTCTCTGT 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 330 TTTAAGAYWTAAGAAGCCCTCCACCTGTTCTGTTTCAAAATAGGCTCTGCAMTTATCCGGG 389
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 958 GTTTTCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 390 AAAGCAGAAAGTGTTCCTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 TGGTTGTTTTCAGATCGATGTCATTTGTGAAGATAGTCCGAGACAGGTGAAGAGCTGA 1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 450 AAACGNTGAGTTCCAGAACTTCAATTTGCAATTTGAGTTCCAAATAGAAATCATGW 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1078 AAAGGCTGAACCTTTGATAGGGAAACTGTTGAGACCTTATTTGCTAAGTGTAAGAAAT 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 510 ATGGTATCTCATATAATAATCCAACTGCTGATAGATTTTCTTAAATGAGAAACAGACAG 569
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 475
; LENGTH: 8709
; TYPE: DNA
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-475

Query Match          1.5%; Score 37.2; DB 7; Length 8709;
Best Local Similarity 56.6%; Pred. No. 13;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 2036 TTGACATAGATGGCGCAATCCCTAACCAACACACACTGATTTCAAGTAAACCTCAGTAG 2095
DB 1228 TTCAACAATACCGGTAATAATCACTATAGAAAAGATGCGAGCTTTAATAACACTTCGTTTC 1287

QY 2096 AACACATCTTTTGACACAACTACTGAAATCAAGATTAGTGGTTTTGATTAGGGAATTTAA 2155
DB 1288 AACACTCTCTGTGTATACAAACACATGACTATTAGTGGTGGGTTTACTTTTAAAGCGGTAAA 1347

QY 2156 AA 2157
DB 1348 AA 1349

RESULT 10
US-11-136-527-7976
; Sequence 7976, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7976
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7976

Query Match          1.5%; Score 36.8; DB 7; Length 600;
Best Local Similarity 59.6%; Pred. No. 3.6;
Matches 62; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2188 CATTGATGAGGGCTAGAAAATGTGTAGGCGATCAGTTTGCATTTGATGGAGGCAATTG 2247
DB 1 CCTTCTCCGCTGGGCCAGGAATGTCATAGGACAGACTTTTGCATGACAGATGAAGG 60

QY 2248 TGGCACTCCGGTGTCTTCACGGGTTAAACGTTGAGCTGGT 2291
DB 61 TGGCACTGGCGTGACGCTGCTGCGCTTCGCGTCTCGCGGAT 104

RESULT 11
US-10-750-185-42915
; Sequence 42915, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 475
; LENGTH: 8709
; TYPE: DNA
; ORGANISM: Helicobacter pylori J99
US-10-751-235-5.rnpbn

; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42915
; LENGTH: 986
; TYPE: DNA
; ORGANISM: Bovine 19866881057042
US-10-750-185-42915

Query Match          1.5%; Score 36.4; DB 6; Length 986;
Best Local Similarity 50.0%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1063 ACAGGTGAAGCTGAAAAGGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTATTGC 1122
DB 768 AGATGGGAAAATGCAAAAAGGGGCAAGATTACTTAAAGAAATCCAATGATTTTCTCAAAAT 827

QY 1123 TAAGTCTAAAGAAATTTGCGAAAGAGAGGCGGAAAGAAATCAATGATGAGGAGTATGTAAA 1182
DB 828 ATCTATTTTCTTTATTTGTGCAAAAATAATGCTGACCGACCAACAGGTAAGCATTTCTACAAA 887

QY 1183 TGATGCTGACCCCAAGTATCTCGGTTTCTTGTGCAAGCAGAGAGGTTTAAACTTT 1242
DB 888 AGATACATGATTTAGCTTTTAGGCATTTTCATGTTTCAATGGTTAAGGAGCTTTAAATTTT 947

QY 1243 TT 1244
DB 948 AT 949

RESULT 12
US-10-750-623-42915
; Sequence 42915, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42915
; LENGTH: 986
; TYPE: DNA
; ORGANISM: Bovine 19866881057042
US-10-750-623-42915

Query Match          1.5%; Score 36.4; DB 6; Length 986;
Best Local Similarity 50.0%; Pred. No. 6.2;
Matches 91; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1063 ACAGGTGAAGCTGAAAAGGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTATTGC 1122
DB 768 AGATGGGAAAATGCAAAAAGGGGCAAGATTACTTAAAGAAATCCAATGATTTTCTCAAAAT 827

QY 1123 TAAGTCTAAAGAAATTTGCGAAAGAGAGGCGGAAAGAAATCAATGATGAGGAGTATGTAAA 1182
DB 828 ATCTATTTTCTTTATTTGTGCAAAAATAATGCTGACCGACCAACAGGTAAGCATTTCTACAAA 887

QY 1183 TGATGCTGACCCCAAGTATCTCGGTTTCTTGTGCAAGCAGAGAGGTTTAAACTTT 1242
DB 888 AGATACATGATTTAGCTTTTAGGCATTTTCATGTTTCAATGGTTAAGGAGCTTTAAATTTT 947
```


This Page Blank (uspto)